AD 683302

Technical Memorandum TM-69-2

JANAIR Report No. 680713

JANAIR VERTICAL CONTACT ANALOG DISPLAY EVALUATION PROGRAM

Accuracy of Altitude, Roll Angle, and Pitch Angle Judgments As a Function of Size of Vertical Contact Analog Display

By

K. D. CROSS and A. C. BITTNER
Systems Integration Division

31 January 1969

This research was accomplished for the JANAIR Program under
ONR Project Orders 8-0079, 8-0080, 8-0081, 8-0082

Naval Research Task NR 213-037

NAVAL MISSILE CENTER

Point Mugu, California

TM-69-2 (U)

109

NAVAL MISSILE CENTER

AN ACTIVITY OF THE NAVAL AIR SYSTEMS COMMAND

C. J. RICKETTS, CAPT USN
Commanding Officer

D. F. SULLIVAN
Technical Director

This report describes work accomplished under AIRTASK RAV-09M001 (A), Vertical Display Evaluation. This research was accomplished for the JANAIR Program under ONR Project Orders 8-0079, 8-0080, 8-0081, 8-0082.

Mr. W. A. Eberspacher, Acting Head, Systems Integration Division; and Mr. R. H. Peterson, Head, Laboratory Department, have reviewed this report for publication.

THIS REPORT HAS BEEN PREPARED PRIMARILY FOR TIMELY PRESENTATION OF INFORMATION. ALTHOUGH CARE HAS BEEN TAKEN IN THE PREPARATION OF THE TECHNICAL MATERIAL PRESENTED, CONCLUSIONS DRAWN ARE NOT NECESSARILY FINAL AND MAY BE SUBJECT TO REVISION.

Technical Memorandum TM-69-2
JANAIR Report No. 680713

Published by . . . . Human Factors Engineering Branch
Systems Integration Division
Laboratory Department
First printing . . . . 235 copies
Security classification . . . . UNCLASSIFIED

NOTICE

Change of Address

Organizations receiving JANAIR Reports on the initial distribution list should confirm correct address. This list is located at the end of the report just prior to the DDC Form 1473. Any change in address or distribution list should be conveyed to the Office of Naval Research, Code 461, Washington, D. C. 20360, ATTN: JANAIR Chairman

Disposition

When this report is no longer needed it may be transmitted to other organizations. Do not return it to the originator or the monitoring office.

Disclaimer

The findings in this report are not be construed as an official Department of Defense or Military Department position unless so designated by other official documents.

FOREWORD

This report presents work which was performed under the Joint Army-Navy Aircraft Instrumentation Research (JANAIR) Program, a research and exploratory development program directed by the United States Navy, Office of Naval Research. Special guidance is provided to the program for the Army Electronics Command, the Naval Air Systems Command, and the Office of Naval Research through an organization known as the JANAIR Working Group. The Working Group is currently composed of representatives from the following offices:

U.S. Navy, Office of Naval Research, Aeronautics, Code 461
Washington, D.C.
Aircraft Instrumentation & Control Program Area

U.S. Navy, Naval Air Systems Command, Washington, D.C.
Avionics Division; Navigation Instrumentation and Display Branch (NAVAIR 5337)

Crew Systems Division; Cockpit/Cabin Requirements and Standards Branch (NAVAIR 5313)

U.S. Army, Army Electronics Command, Avionics Laboratory
Fort Monmouth, New Jersey
Instrumentation Technical Area (AMSEL-VL-I)

The Joint Army-Navy Aircraft Instrumentation Research Program objective is to conduct applied research using analytical and experimental investigations for identifying, defining, and validating advanced concepts which may be applied to future, improved Naval and Army aircraft instrumentation systems. This includes sensing elements, data processors, displays, controls and man/machine interfaces for fixed and rotary wing aircraft for all flight regimes.

JANAIR

JOINT ARMY-NAVY AIRCRAFT INSTRUMENTATION RESEARCH

ACKNOWLEDGMENTS

The authors wish to acknowledge the very significant contributions of several individuals who have been closely associated with the JANAIR Vertical Contact Analog Display Evaluation Project since its inception at the Naval Missile Center. Messrs O. M. Powers and F. R. Cavallero have made singularly valuable contributions in the design and layout of the laboratory facilities and in implementing the many special equipment features that were found necessary for the conduct of this research. Lieutenant Commander R. J. Wherry, Jr., USN, gave generously of his time and talent assisting in the formulation of the experimental designs and the special procedures for data analysis. Messrs D. J. Beaumont, J. N. Roland, and C. W. Abern have made an invaluable contribution through their efforts in maintaining the equipment in good working order. Messrs J. D. Kaplan and P. L. Miller provided valuable assistance in both data collection and tabulation. Last, but certainly not least, Mrs. M. C. Leonard is given special thanks for her patient typing of this manuscript and for her assistance in the many duties surrounding project administration.

TABLE OF CONTENTS

Page

ILLUSTRATIONS

SUMMARY

The purpose of this document is to report the first results from a series of experiments establishing and optimizing the utility of the Vertical Contact Analog Display (VCAD). The VCAD concept is that of a computer-generated pictorial display which provides the observer with visual cues as to aircraft orientation. These cues are analogous to those available under contact flight conditions. The specific objectives of this experiment series were: (1) to establish the relationship between size of VCAD display and judgment accuracy for altitude, pitch, and roll estimates, and (2) to obtain estimates of the accuracy with which these quantities may be judged. Four studies were conducted to measure Ss ability to "maintain" a given standard against a forcing function and three studies were conducted to measure Ss ability to recall and utilize an "internalized standard." Both types of experiments were conducted for each of the three flight related quantities. The experimental designs for each of these experiments were fractional replicates with repeated measures. The analytic procedure consisted of the fitting of "response surface" models and the graphical representation of these models. It was concluded that display size had an effect on the accuracy for each of the three flight quantities, but the nature of the effect depended on both the type of judgment task ("maintenance" vs. recalled "standard") and the magnitude of the standard being judged. Generally the effect was more pronounced for the "maintenance" experiments. The results of the experiments as a unit suggest the optimal display traverses a diagonal visual angle of 18.6 degrees for commercial-rectangular displays (e.g., a commercially designated 11-inch display at a 32-inch viewing distance).

INTRODUCTION

The Vertical Contact Analog Display (VCAD)[1] concept is that of a computer-generated pictorial display which provides the observer with visual cues as to aircraft orientation that are closely analogous to those obtained from an aircraft cockpit under contact flight conditions. The central theme of the VCAD concept is that a digital computer, which maintains "awareness" of aircraft orientation and location through the continuous monitoring of sensor outputs, generates a dynamic, three-dimensional appearing picture of a textured ground and sky plane that corresponds closely with the "real-world" view that would be obtained from the aircraft's current orientation and location in space. It was originally hypothesized that presentation of such a display on a cockpit mounted raster-scan tube would provide a pilot with considerable flight information in a highly integrated format.

The purpose of the JANAIR VCAD Evaluation Program is to evaluate the utility of the VCAD concept for use in a variety of Army and Navy aircraft. Because little systematic research has been conducted previously, the initial phase of this program is necessarily exploratory in nature. Initial efforts have been directed at two rather broad areas of uncertainty. The first stems from the fact that a number of differently configured displays can be designed, all of which can be considered as falling within the general framework of the VCAD concept. There is considerable uncertainty as to which configuration(s) would be most conducive to accurate and rapid information extraction. Although a large number of VCAD characteristics are changeable, those of primary interest during the first phase of the program include: (a) the size, shape, and density of both sky and ground texture, (b) the display viewing angle, (c) the position of texture breakpoints, and (d) the size of CRT on which the VCAD is presented. The experiments reported here deal with the latter of these characteristics, viz., display size.

Uncertainty also exists as to the accuracy with which the information necessary for maintenance of aircraft control can be extracted from the VCAD. It is therefore necessary to obtain reliable predictions of the accuracy with which observers can estimate certain flight quantities from the VCAD.

Thus the objectives of the experiments reported here can be stated as follows: (a) to define the relationship between display size and the accuracy of judging certain flight parameters and (b) to obtain reliable estimates of the absolute accuracy with which these flight parameters can be judged.

- - - - - - - - - - - - - - - - - - -

1 The introduction to this report was written with the assumption that the reader has some familiarity with the characteristics of the General Electric Vertical Contact Analog Display. Those readers not familiar with the display are advised to read the "apparatus" section of this report before proceeding further.

BACKGROUND

Because of the very obvious implications that display size has upon space and weight requirements in he cockpit, it is surprising that researchers have not directed more attention to this variable. Only two studies have been found which have a direct bearing on the relationship between display size and flight proficiency. Wilkerson and Matheny (1961) investigated the effect of "display size" on hovering performance in a Model 47-H1 Bell Helicopter. The "display" in this study was the real world, and "size" was varied by varying the size of opening that was cut in an opaque lining which otherwise masked the view from a helicopter cockpit. Four conditions were investigated: (a) the normal view obtained from the right seat of a Model 47-H1 helicopter, (b) an 8-inch square opening cut in the opaque lining that was located such that the horizon would appear in the center of the vertical dimension when the helicopter was in a hover, (c) an 8-inch square opening located adjacent to and directly below the first, and (d) the lower 8-inch square opening was replaced by an 8-inch high by 16-inch wide rectangular opening, otherwise the arrangement was the same as that described in (c) above. None of these four arrangements were found to exert a significantly different influence on hovering performance. Similarly, Fedderson (1962) compared hovering performance when using an 8-inch Contact Analog Display with performance on a 14-inch display. Performance was measured in the Bell Helicopter Simulation Laboratory using the six-degree-of-freedom dynamic platform developed by Franklin Institute. No statistically reliable differences in performance were found for the two display sizes.

Despite the fact that both of these studies indicate that display size has no measurable effect on flight proficiency, several design features limit the degree to which these findings can be generalized. First, the evaluation of the different display sizes was limited to a hovering task--a highly specialized task with rather unique display requirements. To generalize the conclusion of these studies to other types of aircraft and to other mission segments would appear to be extremely risky.

A related point concerns the nature of the hovering task itself. It is a well known fact that hovering is a difficult and complex task which places large demands on the perceptual-motor skill of the pilot. An over-simplified description of this perceptual-motor task is that the pilot must first perceive a deviation from hover by "decoding" his display, and after completing appropriate information integration and decision making activities, he must introduce corrective response by manipulating his controls (response "encoding"). The point to be made is that although both decoding and encoding errors contribute to the overall task variance, it is probable that encoding is the limiting process in the hovering task. That is, it seems probable that the pilot can detect deviations from hover much more accurately than he can introduce corrections for these deviations. If this is indeed the case, even rather large differences in the accuracy with which different size displays can be decoded would be masked by the even larger variability introduced during the response encoding process.

Other limiting features of these studies can be enumerated. For example, as Fedderson (1962) has pointed out himself, the fact that the performance data on the two different size displays were collected at different times could possibly have affected the result. An even more serious shortcoming was that the performance tasks for the two displays were not the same. The task associated with the 14-inch display required the subjects (_Ss_) to control position, attitude, altitude, rotor rpm, and manifold pressure. In contrast, the task associated with the 8-inch display required _Ss_ to control only position and attitude, a considerably easier task than was required for the 14-inch display.

Whether this confounding of variables resulted in biased conclusions being drawn from these studies is not known. However, it is believed that the probability of such an occurrence is sufficiently high to warrant further investigation of the effect of display size on the accuracy with which the VCAD can be decoded.

Stimulus Configuration: Defined

It is a truism that a thorough understanding of the stimulus configuration in each experimental condition is absolutely essential if the implcations of research results are to be fully understood and if results are to be intelligently generalized to non-laboratory situations. A careful examination of the information parameters which are coded on the VCAD, and the way in which these coded cues change as a function of display size is necessary in order to fully define the stimulus configuration and to understand the manner in which the stimulus configuration changes as a function of display size.

Consider first the information parameters that are necessary if the pilot is to maintain an awareness of his aircraft's orientation in space. An information analysis conducted by the Norden Division of United Aircraft Corporation provides a partial answer to this question (see Williams, 1965; and Williams and Kronhom, 1965). Information requirements were defined for representative missions of three different fixed-wing aircraft, one rotary-wing aircraft, and one VTOL aircraft. The information parameters found to be common to all aircraft and all mission segments investigated are listed in the left-hand column of table 1 and entitled "Flight Parameters."

Considering each of the flight parameters individually, the VCAD was examined to determine what features of the display (if any) change as a direct function of changes in the value of the flight parameter in question. For example, if altitude is increased while all other parameters are held constant, one notes the following changes in the display features: (a) the size of the ground texture elements decreases, (b) the rate at which ground texture elements flow from the horizon to the bottom edge of the display become less, (c) the texture breakpoints

Table 1 Coding of Information on the Basic VCAD

Flight Parameters	Estimates Based Upon
Altitude	Size of ground texture, angular velocity of ground texture relative to current speed, position of texture breakpoints, and angle of convergence of perspective lines formed by ground texture elements
Roll Angle	Inclination of the horizon line and inclination of the perspective lines formed by the sky and ground texture elements.
Pitch Angle	Ratio of visible ground area to visible sky area (when the horizon line falls within the field of view of the display). Position of zenith marker and retinal shape of the sky texture elements (when the horizon is below the display field of view). Retinal shape of ground texture elements and position of texture breakpoints (when the horizon line is above the display field of view).
Airspeed	Angular velocity of ground texture elements relative to current altitude.
Heading Angle	Angle of regard of sky and ground texture elements.
Sideslip	Angular velocity of ground texture elements in a direction perpendicular to the longitudinal axis of the aircraft (relative to current altitude).
Angle of Attack	Ratio of visible ground area to visible sky area relative to the aircraft's velocity vector.

sweep toward the bottom edge of the display, and (d) the angle of convergence of perspective lines decreases, thus causing perspective lines to appear more parallel. Such an analysis serves to identify the population of display features that could possibly cue the observer as to the current value of any one of the flight parameters listed in table 1. The features found to be associated with each of the flight parameters are shown in the right-hand column of table 1. Any estimate of the value of the flight parameters must be obtained through observation of this population of display features.

Summarizing table 1, it can be seen that the population of display features that serve as prospective cues for one or more of the flight parameters inclues: (a) ground texture size, (b) position of texture breakpoints, (c) inclination of horizon line and perspective lines, (d) angular velocity of ground texture, (e) angle of convergence of perspective lines, (f) position of horizon line in vertical plane, (g) retinal shape[2] of ground texture, and (h) angle of regard of sky and ground texture.

Having identified the critical features of the stimulus configuration, the task at hand is to define the effect which modification of display size would have upon each of these eight display features. Where a display feature is found to be measurably different from one display size to another, we ask ourselves, "Would the resulting change lead one to expect that the display feature in question would be more or less discernable, depending upon display size?" An affirmative answer to this question would lead one to predict that the accuracy with which the associated flight parameter could be estimated would be similarly affected if, in fact, that display feature was being attended to. Each of the display features listed above will be discussed in turn.

Ground Texture Size

If display viewing angle[3] is held constant while display size is varied, it can be shown that the linear size of all elements

- -

2 The term "retinal shape" refers to the shape of an image falling on the retina. Retinal shape may or may not correspond with the apparent shape of an object; i.e., the perceived shape. Discrepancies between retinal shapes and apparent shapes provide information as to the perspective from which an object is being viewed.

3 Display viewing angle refers to the angular extent of the world which is displayed on the VCAD. A graphical definition can be found in figure 1 (see angle α). It should also be pointed out here that both viewing distance (distance of VCAD from S) and display viewing angle were held constant in all experiments. Holding these quantities constant while varying display size results in the lack of a one-to-one correspondence between certain display features and the real world. Specifically, the shapes of both sky and ground texture are slightly distorted and the degree of movement of the VCAD horizon line that results from a given increment in pitch differs from display to display and differs from what would be seen if the real world were being observed

appearing on the display varies in direct proporation to display size. However, a description of changes in the stimulus configuration must be expressed in terms of the changes that take place at the eye rather than those which take place at the display. For this reason, the effects of display size upon ground texture size will be discussed in terms of the "visual angle" subtended by the ground texture elements. The visual angle subtended by any object situated perpendicular to the line of sight may be defined as follows:

$$\alpha = 2 \tan^{-1} \frac{h}{2d} \tag{1}$$

where
α = visual angle
h = height (width) of object being viewed
d = distance between the eye and the midpoint of the object being viewed

Examination of a table of trigonometric functions shows that the tangent function is nearly linear over a limited range of values and that $\tan \alpha = \alpha$(in radians) throughout this range. Over the range of values of interest here, 0 to 15 degrees, it can be shown that the maximum error resulting from this approximation is .0061 radians or about one-third degree. Applying this approximation to equation 1, the following simplified expression may be derived:

$$\alpha \approx \frac{h}{d} \text{ (expressed in radians)} \tag{2}$$

It has already been stated that the size of an element of ground texture is directly proportional to the size of display upon which it is presented. Equation 2 shows that the visual angle is, in turn, very nearly proportional to the measured size of that element. Therefore, the visual angle subtended by an element of ground texture can be considered to be a linear function of display size within the range of angles specified above.

The VCAD is constructed such that ground texture elements expand and contract in size as an inverse function of altitude. This expansion and contraction, or more specifically the expansion and contraction of the visual angle subtended by ground texture elements, serves as a "yardstick" by which altitude can be judged. We are concerned here

---

3. Cont.

from a corresponding point in the sky. The distortion in the shape of ground and sky texture could not be detected by any of the Ss used nor were there any detectable differences among the different displays used. The lack of real world correspondence in pitch angle will be discussed in more detail in a later section.

with the question of whether the sensitivity of this "yardstick" might differ depending upon the size of display being viewed. The relationship of interest, then, is the relationship between amount of change in visual angle per increment in altitude on the one hand, and display size on the other.

The amount of change in visual angle resulting from a given change in altitude can be expressed either in terms of relative change or in terms of absolute change. It should be apparent from the above discussion that relative change in visual angle is independent of display size whereas absolute change is linearly related to display size. To illustrate, consider that an element subtending 4-degrees visual angle on one display will subtend 8-degrees visual angle on a display twice as large. An increment in altitude sufficient to reduce texture size by one-half on the smaller display will also reduce texture size by exactly one-half on the larger of the two. Consequently, whereas relative change is seen to be the same for both displays (50 percent), absolute change is twice as great for the larger display (2-degrees versus 4-degrees).

The question that finally emerges is whether the ability to judge change in the size of an object is primarily a function of the absolute or the relative change in the size of an object. Because similar questions concerning the effect of relative and absolute change will arise repeatedly in the following pages, it is believed that this question should be thoroughly considered here.

Classical psychophysical theory (see for example Thurstone, 1927; and Guilford, 1954) states that the accuracy with which change in a stimulus can be detected is a function of the relative change in the stimulus. That is, the value of the stimulus must be changed by some constant fraction of its original value before change will be detected. This psychophysical "law" has been supported using a wide variety of different stimuli and sensory modalities. The most relevant research for present purposes was conducted by Fox, et al. (1959) whose results showed that the area of the squares on a grid has to be changed by a constant fraction (8.1 percent) of its original value in order for subjects to detect a change in size fifty percent of the time (the fifty-percent point served as an arbitrarily selected definition of a "just noticeable difference").

The procedure used in the Fox experiment, and others like it, was the "Method of Paired Comparisons" in which the subject is successively shown a standard and a comparison stimulus and asked to judged whether the two appeared different. It should be noted that because of the time interval separating the two stimulus presentations, the task necessarily involves the memory capacity of the observer--a critical characteristic of the procedure. In the truest sense of the word this procedure requires the observer to <u>judge</u> whether or not a change has occurred rather than to

detect change as is usually implied. The point being made here is that the above-mentioned psychophysical "law" should be generalized to only that portion of a pilot's task that calls upon the same type abilities that are tapped by the Method of Paired Comparisons, viz., tasks in which a pilot is required to compare the current value of a stimulus with his memory of a "standard" stimulus value.

Consideration shows that pilots are indeed confronted with analogous tasks when controlling an aircraft. Any time a pilot is required to judge the absolute value of one of the flight parameters from the VCAD, he is required to utilize his memory capacity in a fashion analogous to that described above. In making absolute judgments, the pilot is trying to assign a numerical value to one or more of the flight parameters using only the VCAD as reference. Assume, for the sake of illustration, that a pilot desires to level his aircraft at one-thousand feet altitude. Such a task requires the pilot to continuously compare his present display configuration with his memory of what the display should look like at one-thousand feet altitude.

Could the pilot perform such a task more accurately on a larger display? It has been shown that the ability to judge whether a current stimulus is different from a recalled standard is a function of the relative difference between the two. It has also been shown that relative change in size per altitude increment is the same for all size displays. It therefore seems reasonable to assume that those tasks involving absolute judgments of flight parameters, such as judging altitude from the size of ground texture, would be made with about the same degree of accuracy regardless of the size of display being utilized.

There are, however, tasks which a pilot must perform other than those involving absolute judgments of flight parameters. For example, pilots are frequently concerned with maintaining a given flight orientation over an extended period of time. Such a task requires that the pilot continuously monitor the VCAD in an attempt to <u>detect</u> and null a change in one or more of the flight parameters. Obviously the ability to detect change is a function of such variables as the rate at which the change takes place and pilot vigilance as well as the magnitude of the change, but it is only the latter variable that is of interest here. Specifically, we are concerned with whether detection of change under conditions of continuous viewing is a function of relative or of absolute change in the stimulus.

Whereas judgment of a change (as previously defined) is dependent upon memory capacity, <u>detection</u> of a change would appear to be more a function of perceptual capacity or, more precisely, visual acuity. Although no research evidence has been discovered which relates detection of size changes with magnitude of the change, it is known that the retina is quite sensitive to changes in position. If detection of a change in the size of

an object is thought of as detecting a change in the position of one or more of its sides, it seems intuitively obvious that the initial size of the object being viewed would have little effect on detection ability so long as at least one of the sides fell within the observer's field of view. Thus, the most feasible guess is that one or more of the edges of an object must traverse some constant linear distance in order for a change in size to be detected--regardless of the initial size of the object being viewed.

Considering that the absolute change in the size of ground texture elements resulting from a given increment in altitude is greater for larger displays and given the assumption that the size of the threshold for change detection is some constant linear value, it follows that a just noticeable change in size would represent a smaller change in altitude for larger displays. Thus, the ability to detect changes in altitude based upon size of ground texture elements should improve as display size is increased.

In an attempt to further clarify this rather difficult section, the following points are reiterated:

(a) The size of the ground texture elements is inversely proportional to altitude and thus serves as one possible cue to altitude and altitude change.

(b) Holding altitude constant, the size of the ground texture elements is directly proportional to the size of display being utilized.

(c) The <u>relative change</u> in size of the ground texture elements resulting from a given increment in altitude is the same for all size displays.

(d) The <u>absolute change</u> in size of the ground texture elements resulting from a given increment in altitude is linearly related to display size.

(e) It is predicted that the ability to make <u>absolute judgments</u> of altitude, based upon the cue of ground texture size, is independent of display size.

(f) It is predicted that the ability to <u>detect</u> changes in altitude based upon size of ground texture elements should improve as display size is increased.

<u>Position of Texture Breakpoints</u>

Texture breakpoints are discussed in detail in the "apparatus" section of this paper and the reader unfamiliar with this feature of the VCAD is strongly advised to refer to this section before proceeding further. The content of this discussion, however, is briefly summarized below.

The ground plane texture is composed of a hierarchy of four nested orders of texture elements. Whereas the texture elements within any given order are all of the same size and shape, the element size and shape differs considerably from order to order. The unique shapes of the texture elements within each of the four orders in the hierarchy are shown in figure 9. The dimensions of the texture elements increase by a factor of eight at each step up the hierarchy--with the result that the fourth order elements are 512 times as large as first order elements. The CADG is constructed such that a texture element within a given order is not defined (not displayed) when the distance from the aircraft to that element exceeds some preestablished distance. Thus, for example, first order texture elements may be seen to extend up to 500 yards into the distance. However, at this point on the earth plane, these elements abruptly terminate and only second and higher order elements can be seen at distances beyond 500 yards. This line of demarcation between orders of texture is what is referred to as a breakpoint. The first, second, and third orders of texture have breakpoints but the fourth order texture extends to infinity.

Although not a deliberately designed visual cue, these breakpoints have been found to provide powerful information for judging altitude and, under certain limited conditions, pitch. Because the breakpoints always occur at a fixed slant range, the breakpoints move toward the bottom of the display as altitude is increased, and toward the horizon line when altitude is decreased. When altitude is held constant, the breakpoint maintains a constant poisition relative to the horizon line regardless of the orientation of the aircraft. Results to date have shown the breakpoints to be a powerful and frequently used altitude cue.

Having defined the manner in which breakpoints serve as cues, attention can be returned to the question of primary concern, viz., whether display size would be expected to affect the powerfulness of the texture breakpoint cue. It was mentioned in the previous section that if display viewing angle is held constant while display size is varied, the linear size of elements appearing on the display varies in direct proportion to display size. Similarly, it can be shown that the absolute magnitude of the change in breakpoint position which results from a given increment of altitude is proportionately greater for larger displays, whereas, relative change remains the same for all size displays. If an increment in altitude is sufficient to move the breakpoint position by a distance of one inch on the face of a 5-inch display, the same increment in altitude would move the position of the breakpoint the same relative amount--two inches on a 10-inch display.

It should be noted that the nature of the relationship between amount of breakpoint change per increment in altitude and display size is precisely the same as that described for size of ground texture. For this reason, the predicted effects of display size on the accuracy of altitude judgments based upon breakpoint position are the same as those presented in the

previous section. Specifically, it is predicted that (a) the accuracy of absolute judgments of altitude based upon the breakpoint cue should be unaffected by display size and (b) the ability to detect altitude changes from the breakpoint cue should improve as display size increases.

Inclination of Horizon Line and Perspective Lines

The most familiar cue for roll angle is the degree of inclination of the horizon line. However, it should be realized that roll angle information is also provided on the VCAD by the degree of inclination of the perspective lines formed by the ground texture elements and by those formed by the sky texture elements. Thus, an observer can learn to associate roll angle value with inclination of perspective lines in the same way that he establishes such associations with the horizon line.

Modifying display size has no effect on the angular inclination of the horizon line and perspective lines for there is always direct correspondence with real world inclination, regardless of the size of display being used. The sole effect of modifying display size--so far as roll angle cues are concerned--is to proportionately modify the length of the horizon line and perspective lines. This fact is clearly illustrated in figure 1 which shows identical inclinations on two different size displays. In speculating as to whether judgments of the inclination of a line would be effected by the length of the line being judged, one point becomes immediately obvious. Increasing the length of a line rotating about a common point increases the linear excursion through which the end of this line travels for a given angular change; therefore, if subjects attend to that portion of the horizon line furthest from the central axis, and preliminary evidence suggests that they do, then larger displays may provide a more sensitive index of change in roll angle.

Before any specific predictions can be set forth, it is necessary to carefully define the nature of the relationship between linear movement of the tip of the horizon line and display size (this discussion is equally valid for perspective lines). This relationship is made somewhat more complicated by the use of rectangular displays. Reference to figure 1 will show that the length of the horizon line is affected not only by display size but also changes as a function of the degree of horizon line inclination (α). As inclination is increased from 0 to 90 degrees, the length of the horizon line increases to a maximum at 39 degrees--the point at which the horizon line intersects with the corner of the standard, rectangularly shaped raster-scan display--then decreases to a minimum at 90 degrees. It should also be noted that the horizon line displacement is in the vertical plane when the absolute value of α is less than 39 degrees and in the horizontal plane when α is greater than 39 degrees. Thus, a function which adequately expresses the relationship between linear displacement of the horizon line and display size must take into consideration display size, the angle of inclination, the amount of change in inclination,

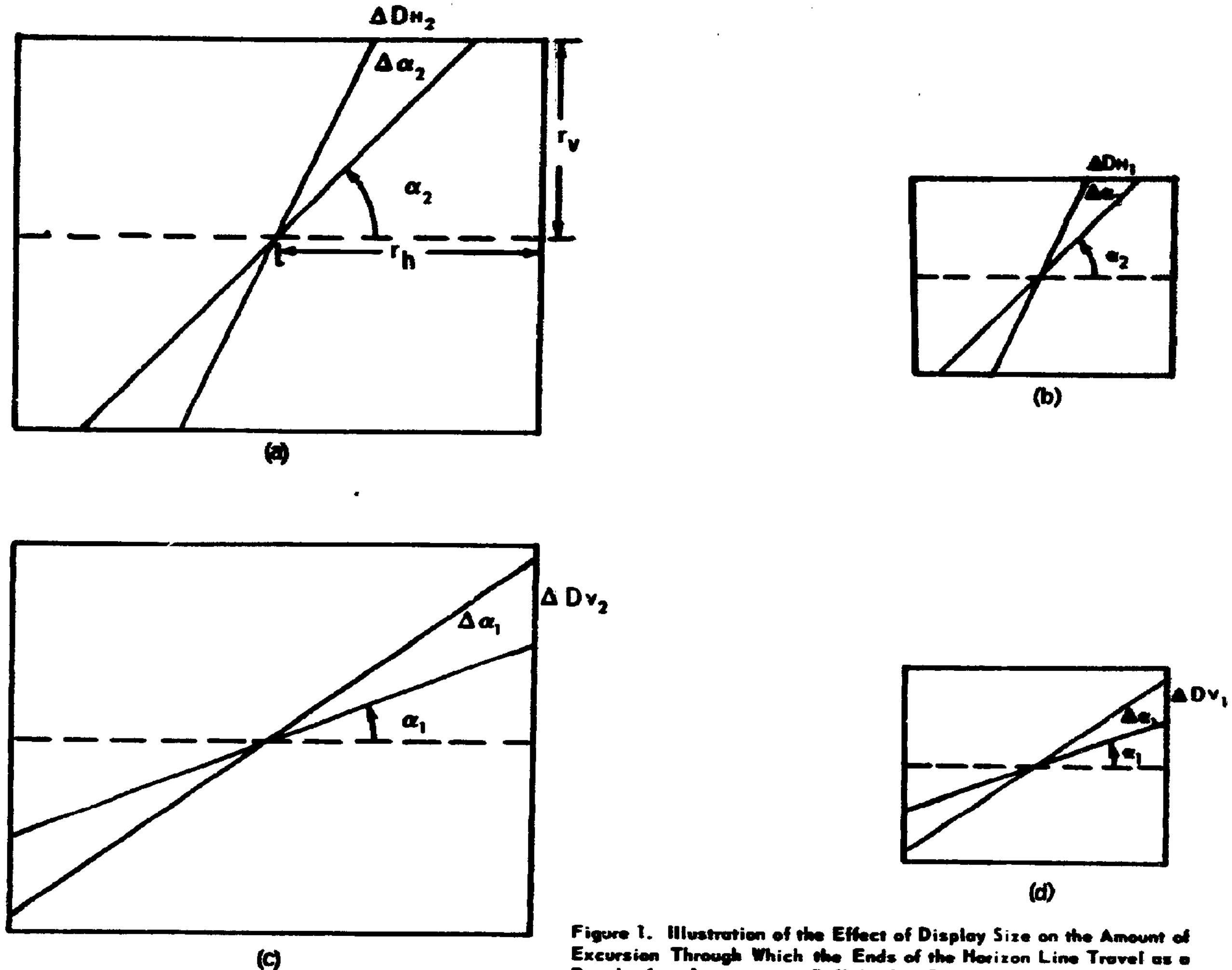


Figure 1. Illustration of the Effect of Display Size on the Amount of Excursion Through Which the Ends of the Horizon Line Travel as a Result of an Increment in Roll Angle. Drawings (b) and (d) are one-half the size of (a) and (c).

and whether the displacement occurs in the vertical or horizontal dimension of the display. The following two equations were derived in accordance with these requirements:

$$\Delta Dv = \frac{r_v \sec^2 \alpha}{\text{ctn}\,\Delta\alpha - \tan\alpha}, \text{ for: } 0 \leq \alpha \geq 39° \tag{3}$$

$$\Delta Dh = \frac{r_h \csc^2 \alpha}{\text{ctn}\,\Delta\alpha + \text{ctn}\,\alpha}, \text{ for: } 39° \leq \alpha \geq 90° \tag{4}$$

Equation 3 defines displacement of the horizon line in the vertical dimension and is to be used when the absolute value of α is less than 39 degrees. Equation 4 defines displacement in the horizontal dimension and is to be used when the absolute value of α is between 39 and 90 degrees. These equations express displacement as a function of display size (r), inclination (α), and increment in inclination ($\Delta\alpha$). These quantities are defined graphically in figure 2.

The equations clearly illustrate that the excursion through which the tip of the horizon line travels (ΔD) as a result of a given increment in inclination is a linear function of display size. Although less obvious, it can also be shown that when all quantities are held constant, ΔD approaches a maximum as α approaches 39 degrees.

It should be clear that when α approaches 39 degrees and/or when display size is increased a given linear excursion of the horizon line along the edge of the display represents a smaller angular change. Following the rationale presented earlier, it is predicted that the ability to detect change in roll angle would increase as a function of display size. Furthermore, for a given size display, it is predicted that ability to detect change would increase as roll angle approaches 39 degrees. Because relative change is independent of display size, it is predicted display size would have no effect on the accuracy with which absolute judgments of roll angle could be made.

Angular Velocity of Ground Texture Elements

In a contact flight situation the angular rate at which a point on the ground sweeps beneath the aircraft is a function of: (a) the line of sight to the ground point, (b) the aircraft altitude, and (c) aircraft speed. Therefore, the angular velocity of elements on the ground plane serves as a prospective cue for judging altitude and/or speed. The relationship between the angular velocity of any point on the ground and the altitude and speed of the aircraft is expressed by the following function (for a contact flight situation):

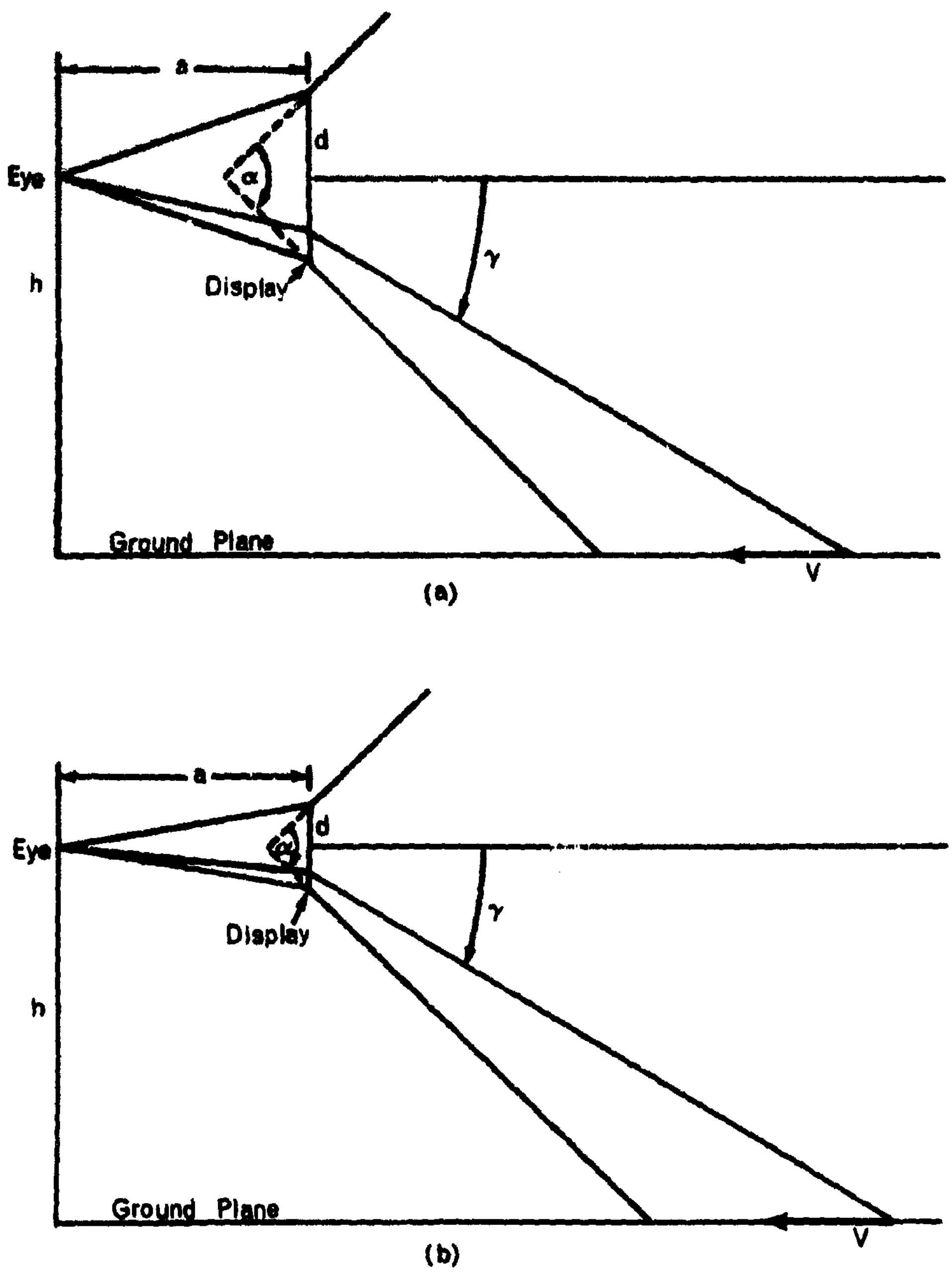


Figure 2. Graphical Definition of the Terms Used in Equation (4). Equation (6) defines angular velocity of ground texture elements as a function of six relevant variables.

$$\omega = \frac{V \sin\alpha}{h} [\cos^2 \beta \sin^2 \alpha + \sin^2 \beta]^{1/2} \tag{5}$$

Where: ω = angular velocity at eye of observer
V = ground plane velocity
α = lookdown angle measured from longitudinal axis of aircraft
β = lateral look angle measured from longitudinal axis of aircraft
h = altitude

It is apparent from equation (5) that the angular velocity of a given point on the ground is directly proportional to ground plane velocity and inversely proportional to the altitude. The stimulus configuration presented on the VCAD is directly analogous to the stimulus configuration obtained in contact flight if and only if display size, display viewing angle, and viewing distance (distance from eye of observer to center of display) are such as to produce a one-to-one correspondence with the real world. In order to define the stimulus configuration for those situations in which the display does not correspond with the real world, the following expression was derived. The terms in equation (6) are also defined graphically in figure 2.

$$\omega = \frac{Vad \operatorname{ctn}(\frac{\alpha}{2})}{h [a^2 \operatorname{ctn}^2 \gamma + d^2 \operatorname{ctn}^2 (\frac{\alpha}{2})]} \tag{6}$$

Where: ω = angular velocity at eye of observer (in radians)
V = ground plane velocity
a = viewing distance (eye to display)
d = one-half display height
α = display viewing angle
h = altitude
γ = lookdown angle = $\beta + \theta$ Where: β = positive angle down from horizon to line of sight to ground point

θ = negative angle down from horizon to center line of display.[4]

The persisting theme throughout this section has been the illustration of the relationship between the powerfulness of a displayed cue and the size of display being utilized. Unfortunately the relationship between angular velocity and display size, which is the relationship of interest here, is not intuitively obvious from equation 6. In order to clarify this relationship, the angular velocity of a point on the ground plane has been plotted as a function of display size for two altitudes and two speeds (see figure 3). Viewing distance, display viewing angle, and

---

4. When pitch angle is zero, then $\theta = 0$ and $\gamma = \beta$. The illustrations in figure 2 were drawn with a zero pitch angle; therefore, since $\gamma = \beta$ only γ is shown.

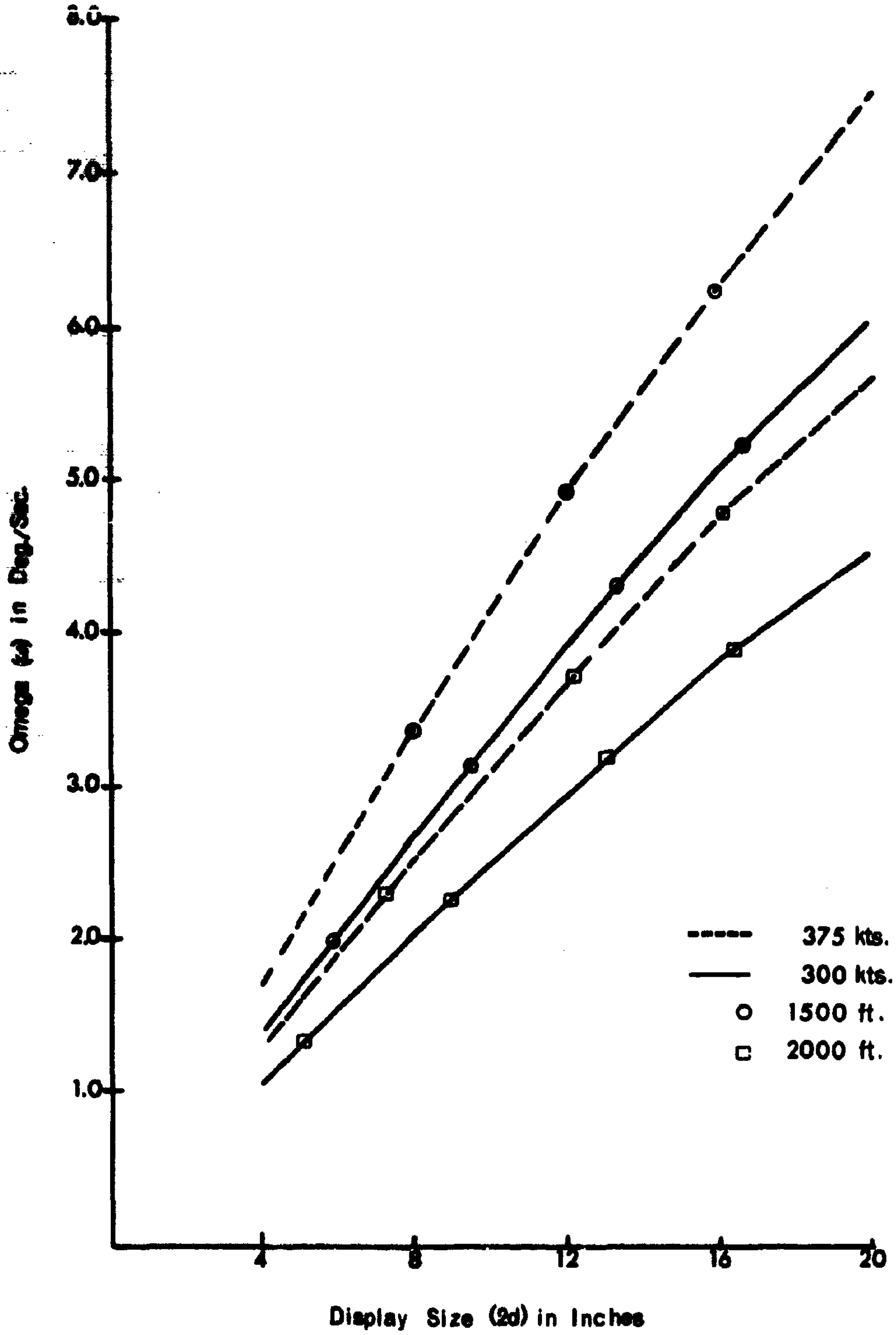


Figure 3. Angular Velocity of Line of Sight to a Moving Point on the Display as a Function of Display Size, Altitude, and Actual Ground Plane Velocity.

lookdown angle were held constant at 26-inches, 90-degrees, and 40-degrees respectively.

Although not the relationship of primary interest it should be noted in figure 3 that angular velocity is directly proportional to both display size and horizontal velocity but inversely related to altitude. Furthermore, these relationships are very nearly linear. Of more immediate interest is the size of the increment in angular velocity that results from an increment in horizontal velocity or from a decrement in altitude, remembering that it is this increment that is the yardstick by which changes in speed and altitude must be judged.

First examine, for each display size, the increment in angular velocity that results from a decrement in altitude of 500 feet. This increment is represented by the vertical distance between the two solid lines (V = 360 kts) and between the two dashed lines (V = 375 kts). The fact that these curves systematically diverge as display size is increased clearly illustrates that the magnitude of the change (decrement) in angular velocity is greater for larger displays. It would therefore appear that a given altitude change would be more likely to be detected on a larger display.

The increment in angular velocity resulting from a 75 kt/hr increment in horizontal velocity is represented by the vertical separation between the pairs of curves coded with the same symbol. Once again it is seen that the curves diverge as display size increases--showing that a given increment in horizontal velocity results in a larger increment in angular velocity when a larger display is used. The conclusion is similar to that presented above; viz., speed changes should be more accurately detected on larger displays.

Angle of Convergence of Perspective Lines

The ground plane may be thought of as a matrix of essentially infinite extent. What has been referred to here-to-fore as a ground texture element resides within each cell of the matrix, giving the ground plane the appearance of a completely symmetrical series of rows and columns of texture elements on a homogeneous background. Unless pitch angle equals ± 90 degrees, the columns of texture elements form perspective lines which converge at the horizon. At very high altitudes these perspective lines appear very nearly parallel, whereas at very low altitudes the angle of convergence approaches 180 degrees. The systematic change in degree of convergence that results from altitude change provides for the realistic, three-dimensional appearance of the display and also provides a powerful cue to altitude. A quantitative association between degrees of convergence of perspective lines and altitude can be established in one of three ways. These are illustrated in figure 4 which shows the manner in which a set of perspective lines change as a result of an increment in altitude. These changes are illustrated for two different size displays, one superimposed on the other. The solid lines represent the original position of

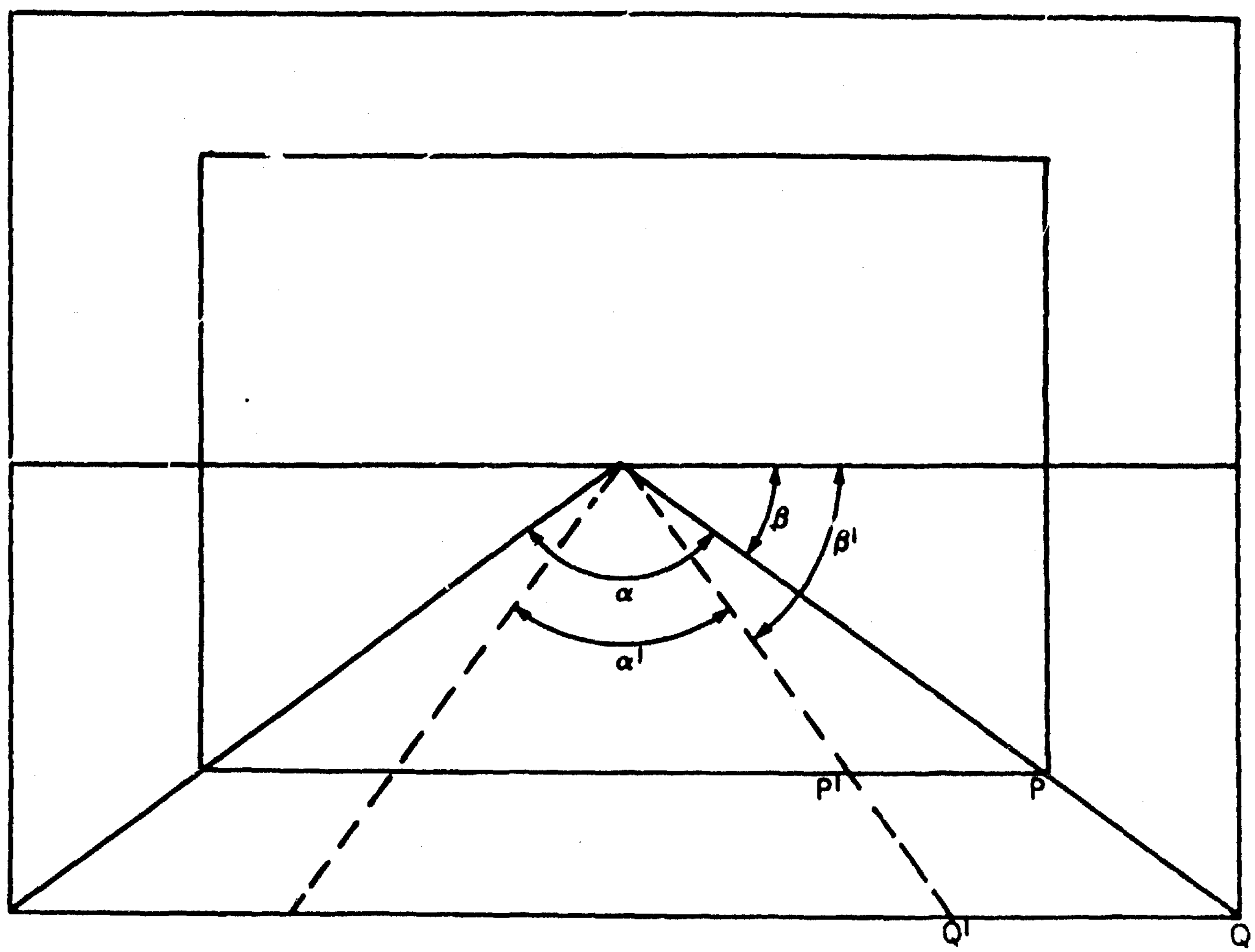


Figure 4. Illustration of the Effect of Display Size on the Altitude Cues Provided by the Perspective Lines. The outer rectangle represents the perimeter of a "large" display and the smaller rectangle represents the perimeter of a display half as large. The dotted lines show the position of the perspective lines after an increment in altitude.

the perspective lines and the dashed lines represent the position after an increment in altitude has been introduced.

The observer can learn to associate altitude with: the angle formed by two perspective lines (see angle α in figure 4), the angle formed by a perspective line and the horizon line (see angle β in figure 4), or the point at which a perspective line intersects the edge of the screen (see point P in figure 4). All of these quantities are sensitive to altitude change and, with training, an observer can learn to associate a given value or change in value with a given altitude or altitude change.

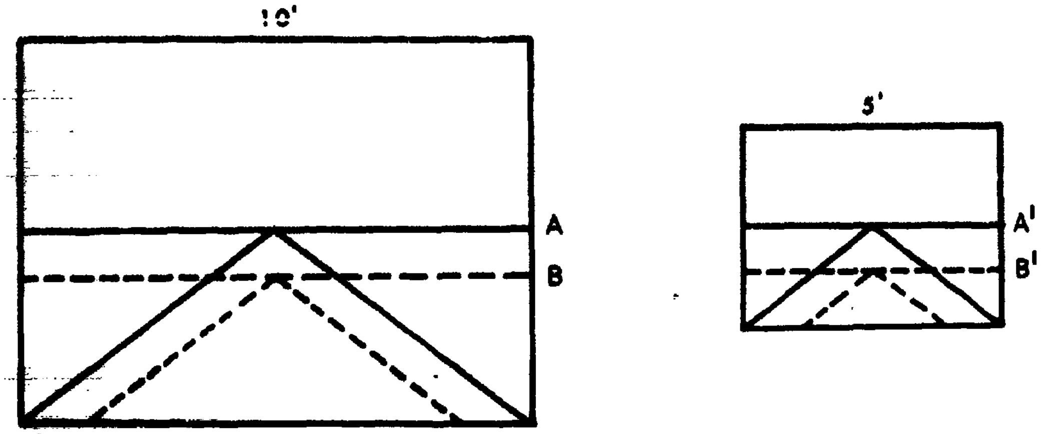
The amount of change in angle α and β per increment in altitude is the same regardless of the size of display being used. However, the linear excursion through which a perspective line travels as a function of an increment in altitude is greater for larger displays. The differences in the length of lines PP' and QQ' illustrate the increase in linear excursion obtained on a larger display. Therefore, if the observer chooses to associate the difference between angles α and α' or between β and β' with altitude change, it is predicted that display size would have no effect on judgment accuracy. But if the observer references the point at which a perspective line intersects the edge of the screen and associates change in position of this point with change in altitude, it would be predicted that larger displays would facilitate altitude judgment accuracy.

Position of Horizon Line in Vertical Plane

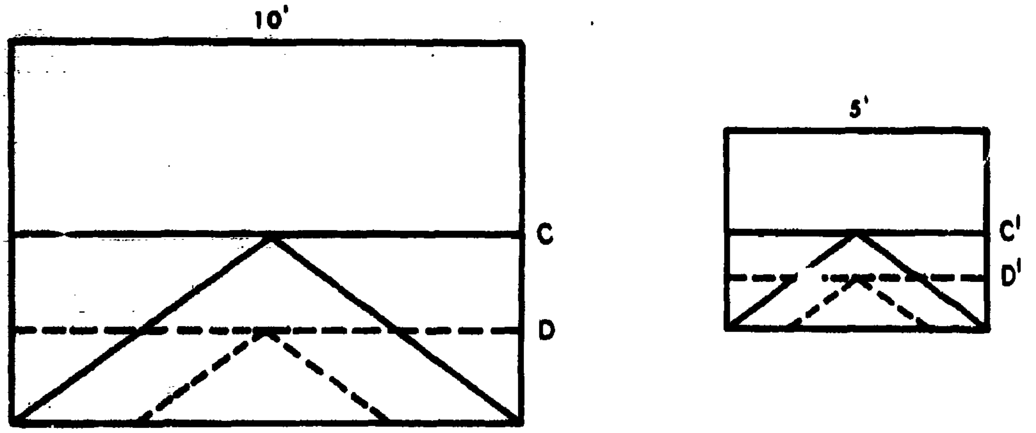
The position of the real world horizon line is the most direct and familiar indicant of pitch angle in contact flight. A 15-degree change in the position of the horizon line (in terms of visual angle at the eye) indicates a 15-degree change in pitch angle. The VCAD horizon line is also an indicant of pitch angle but the amount of position change per increment in pitch angle may or may not correspond to changes in the real world horizon line. The amount of change in position of the displayed horizon line is a function of not only pitch angle but display size and display viewing angle as well.

The relationship among these three variables is illustrated in figure 5. The larger line drawings are twice the size of the smaller. The solid lines represent the position of the horizon line and a set of perspective lines at zero pitch angle, whereas, the dashed lines represent their position after a 15-degree increment in pitch.

The set of drawings marked (a) shows the nature of the stimulus configuration when the display viewing angle is manipulated such as to produce a one-to-one real world correspondence on both size displays. The equality of lines AB and A'B' shows that the amount of horizon line displacement per pitch increment is the same for both displays. When display viewing angle is held constant as display size is varied--as



(a) 15° Pitch Increment When 1 to 1 Real World Correspondence is Maintained.



(b) 15° Pitch Increment When 60°-Display Viewing Angle is Maintained.

Figure 5. Effect of Display Size and Display Viewing Angle on the Amount of Displacement of the Horizon Line Resulting From a 15° Increment in Pitch.

shown in the set of drawings marked (b)—the amount of horizon displacement per pitch increment increases as a function of display size. The large display is twice as large as the smaller, thus the amount of horizon displacement is twice as great on the larger display; i.e., CD = 2C'D'.

It can thus be seen that larger displays provide a more sensitive index of pitch angle if display viewing angle is held constant but that no such advantage exists if a one-to-one correspondence with the real world is maintained across all size displays. Assuming display viewing angle is held constant, it is predicted that increasing display size would facilitate detection of pitch changes.

Shape of Ground and Sky Texture Elements

When positive or negative pitch angle exceeds one-half the display viewing angle, the horizon line disappears from the display's field of view and is no longer available as a pitch reference. In this situation the only stable cue for pitch angle--other than the "Zenith Marker" which will be discussed later--that is independent of the other flight parameters is the shape of the ground and sky texture elements. It is only when pitch is equal to positive or negative 90 degrees that the "retinal shape" of the texture elements (the shape of image falling on the retina) is the same as the "apparent shape" of the elements (the perceived shape or the shape the elements are known to be). At all other pitch angles a perspective view of the texture elements is obtained such that the retinal shape is different from the apparent shape. That is, the retinal shape of square objects is actually trapezoidal and the retinal shape of round objects is ovoid. Despite the phenomenon of shape constancy (objects are perceived as being the same shape from all angles of regard) a human observer is quite sensitive to discrepancies between apparent shape and retinal shape. It is this difference that serves as a cue to pitch angle.

Although no evidence is available to suggest whether sensitivity to this discrepancy is depedent upon the absolute size of the object being judged, it seems highly doubtful that such an effect would be obtained over the range of interest here. It is therefore predicted that display size would have no appreciable affect on the judged discrepancy between retinal and apparent shape of texture elements.

Angle of Regard of Sky and Ground Texture

Sky and ground texture elements maintain a fixed orientation on the displayed sky and ground plane. As heading is changed, these elements are viewed from a correspondingly different angle of regard. Ability to judge heading angle is therefore a direct function of the observer's ability to judge the angle from which an element is being viewed. As the size of elements on even the smallest display being considered is well above the acuity threshold, and since the ability to judge angle of

regard would appear to be independent of the size of object being viewed so long as it can be clearly seen, it seems doubtful that display size would significantly affect judgment accuracy.

Overview

In defining the display configuration in terms of the specific display features that serve as prospective cues for one or more of the flight parameters, we have essentially identified a number of "perceptual yardsticks" by which the current value of each of the flight parameters must be judged. Each of these yardsticks was examined analytically in an attempt to predict whether the sensitivity of the yardstick would be affected by changes in display size. The focus of this examination was the amount of change in the display feature that results from a given change in the associated flight parameter. The finding for the majority of the display features examined was the same, viz., the amount of "absolute change" was found to be greater for larger displays whereas "relative change" was found to be the same for all size displays. These findings led to the general predictions that (a) the accuracy with which numerical values could be associated with each of the display features (absolute judgments) would be independent of display size and (b) an observer's ability to detect changes in the flight parameters would improve as display size is increased.

Research Objectives

The experiments reported here were directed at two broad questions, viz., "What is the effect of display size on an observer's ability to extract flight relevant information from the VCAD? and With what degree of accuracy can flight relevant information be extracted from the VCAD?" Thus, the primary independent variable of interest in these experiments is display size and the dependent variable of interest is decoding accuracy. The objectives of the experiments, stated as briefly as possible, is to quantitatively define the relationship between the independent (display size) and dependent (decoding accuracy) variable. These objectives are further clarified in the discussion below.

The rationale for designing experiments to define <u>relationships</u> rather than to simply detect statistically reliable <u>effects</u> should first be clarified. This rationale evolved from a need to generalize beyond the specific display sizes utilized in this study. It was reasoned that if display size did prove to exert a significant influence on decoding accuracy it would be desirable to know what size display would prove optimum. Realizing that there was no assurance that one of the display sizes tested would be the optimum size, it became apparent that the experimental data collected should provide for the prediction of the optimum display size. It was also reasoned that cockpit design considerations might conflict with the utilization of the optimum display size. Obviously, an intelligent trade-off between design efficiency/economy and decoding accuracy can be made only if one is aware of the degree to which decoding

accuracy is being sacrificed as a result of utilizing a non-optimum display size. The above needs can be served only by a reliable estimate of the relationship between display size and decoding accuracy over a broad continuum of display sizes. A final advantage of this approach is that knowledge of the nature of the relationship between two variables often leads to important insights into the basic perceptual process involved--insights which do not frequently evolve from the comparative statistical approach that is traditionally used in hypothesis testing.

The process of interest during these initial experiments is limited to the perceptual process involved in extracting information about various flight parameters from the VCAD. This accounts for use of the term "decoding accuracy" to describe the independent variable rather than a term such as flight proficiency. These experiments were designed to measure decoding accuracy as independently as possible from the other processes normally required in an actual flight situation such as information integration, decision making, response encoding, etc. To accomplish this purpose, task loading and complexity of response encoding were deliberately reduced to a minimum. Obviously decoding ability under such simplified conditions will be superior to that expected in an actual flight situation, but this approach is believed necessary for two reasons. First, if it is found that the VCAD cannot be decoded with sufficient accuracy under these simplified conditions, it will be apparent that some type of supplemental information will be necessary in order to meet minimum information requirements in a flight situation. Second, if it is found that decoding is accomplished with sufficient accuracy under simplified conditions, but breaks down when the control task is made more complex, it will be known that the decoding process is not the weak link in the total perceptual-motor process and attention can be directed at other processes.

Two types of decoding are investigated. One type has been previously identified as that of making absolute judgments and is essentially a measure of how well an individual can associate real world numerical values to the various flight parameters using only the basic VCAD as reference. A second type of decoding has been previously identified as that of detecting change and is a measure of how well an individual can detect a change in the real world value of the various flight parameters--whether or not he can associate a numerical value with this change.

GENERAL METHODOLOGY

The experiments discussed in this report have a great many features in common as far as methodology is concerned. In order to avoid undue redundancy a general discussion of methodology will be presented which is relevant for all the experiments. Then, additional methodological

features which are specific to individual experiments will be discussed where appropriate.

Apparatus

The experiments reported herein were conducted in the Human Factors Engineering Advanced Displays/Controls Evaluation Laboratory. This laboratory was designed to facilitate the conduct of a wide variety of research and thus consists of a rather large number of interconnected and highly versatile components. Two different views of the laboratory are shown in figure 6. A detailed description of each of the components comprising the experimental apparatus would be prohibitively lengthly and would add little to the present discussion. For this reason, the apparatus has been divided into three functional components for discussion.

VCAD Generator

The VCAD Generator is a special purpose digital computation system that accepts position and attitude data from an aircraft simulation system, integrates this data and performs the computations necessary to synthesize a rapidly updated, analogous representation of the aircraft environment as it would appear from an aircraft cockpit. This one-of-a-kind device is an experimental tool which was developed for the sole purpose of evaluating the VCAD concept. Its construction is such that many of the features of the generated display can be conveniently varied over a wide range of value. The important features of the VCAD Generator are discussed below. The reader is referred to volumes A and B of the VCAD instruction manual (General Electric, May 1966) if a more technical discussion is desired.

Sky and Ground Plane: The VCAD Generator defines both a sky plane and a ground plane which appear on the display as planes of infinite extent. The ground plane is computed to be tangent to the earth's surface at the nadir of the simulated aircraft. The sky plane is computed to be parallel to the ground plane and is situated at a fixed altitude above the ground plane. Computationally, the sky and ground planes are composed of an infinite number of square matrices situated adjacent to one another in symmetrical rows and columns. Each matrix, in turn, consists of eight rows and columns of "cells." The VCAD Generator is designed to provide for independent on-off control of the 64 cells within a matrix, but does not provide for independent control of the matrices themselves. That is, although the on-off state of a given cell may be arbitrarily selected, this same state will exist in the corresponding cell of every matrix appearing on the plane.

The above design characteristic defines both the capability and limitation for producing a variety of texture patterns on the sky and ground planes. A homogeneous surface is obtained if all 64 cells are placed in the same state. A textured surface is obtained so long as the

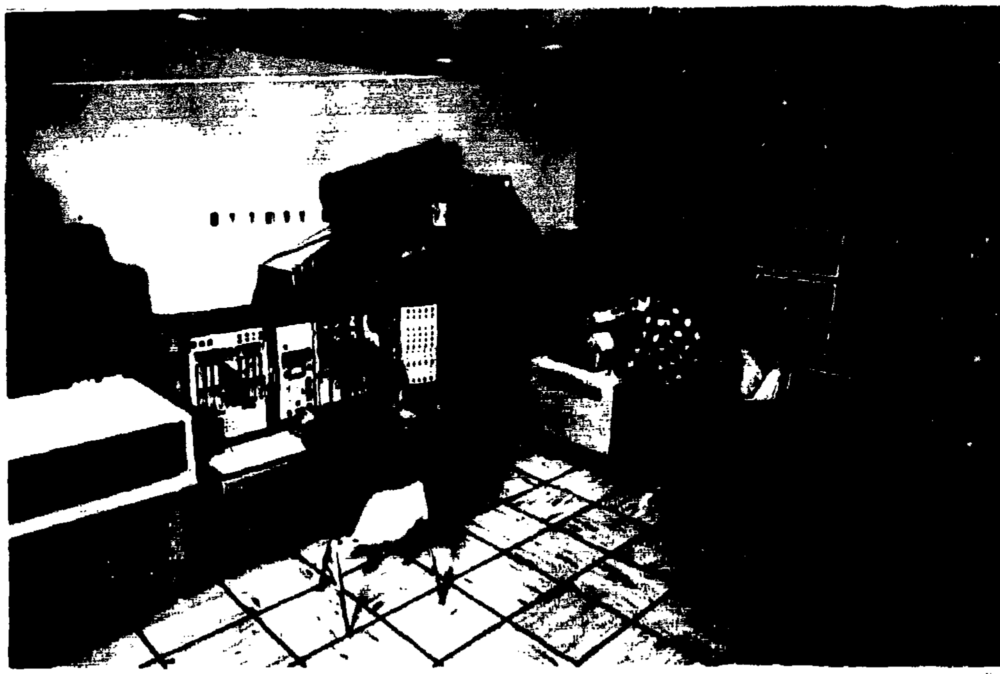


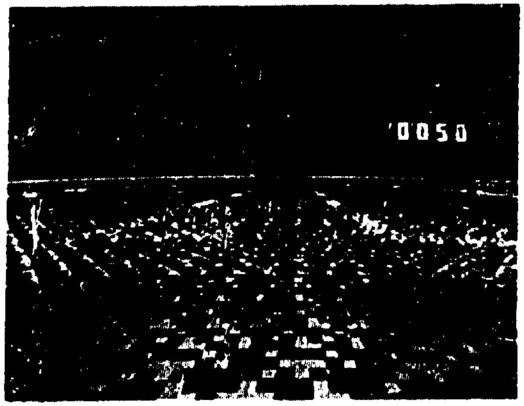
Figure 6. Two Views of the Human Factors Engineering Advanced Displays/Controls Evaluation Laboratory.

state of one or more cells is different from those remaining. The variety of different texture patterns that can be produced is equal to the number of on-off combinations of 64 objects taken one or more at a time. It should be kept in mind, however, that even though a texture was composed by determining the state of each of the cells randomly, the texture of the plane would still appear symmetrical due to the repetition of the pattern in all matrices appearing on the plane.

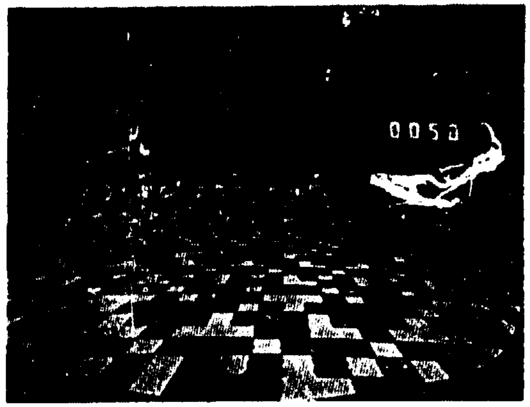
Based upon the parameters which define the momentary position and attitude of the simulated aircraft, the VCAD performs the computations necessary to display an image of the sky and ground plane as they would appear from a similarly situated aircraft. These computations are based upon "real-world" geometric and perspective rules making the view of the sky and ground planes obtained on the display a faithful analogy of real world conditions. (This faithful analogy between the features of a synthetic world and those of the real world enable an observer to utilize those perceptual skills developed during a lifetime of everyday experience and represents the fundamental theme of the VCAD concept.) The computations necessary to present an updated view of the sky and ground plane occur 30 times per second. Although these changes occur discretely, the update rate is sufficiently rapid to produce the effect of continuous change. With one exception, the change in the view of the sky and ground planes which results from a change in one or more of the flight parameters follows real world geometric and perspective rules. This exception is that changes in the sky texture elements are independent of changes in both velocity and altitude. Elements on the sky plane always appear to be the same distance above the simulated aircraft regardless of altitude. Furthermore, the sky texture elements do not appear to sweep overhead as one would expect; rather, their position is fixed with respect to the aircraft regardless of its position and velocity. Thus, the sky plane can be said to respond to rotation about the three axes of the aircraft but does not respond to aircraft translation.

The scaling in all computations is determined by the size of the "cell"-the basic element of both the sky and ground planes. Cell size is variable and can be made to equal 4-, 8-, 16-, 32-, or 64-feet. The effect of cell size upon the display configuration is illustrated in figure 7. The result of changing cell size is obviously to change the size of the ground texture elements on the ground plane. One setting may produce a display configuration that would be analogous to flying over terrain dotted with houses. Another setting may produce a view analogous to flying over a rural area where the terrain is divided into square-mile segments.

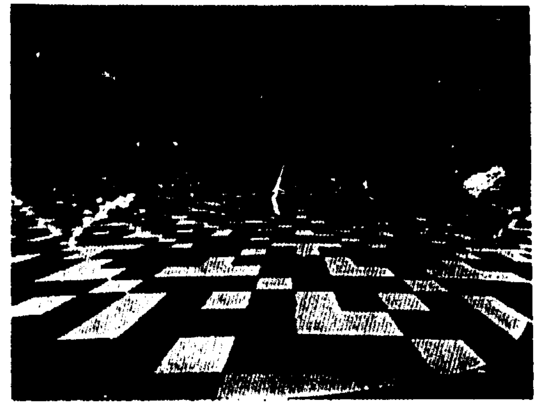
It should also be noted that the size of the basic texture element is also affected by the number of cells used to construct the texture element. This aspect is illustrated in figure 8 which illustrates the manner in which three types of texture elements were constructed on the 8x8 cell matrices and shows their respective appearance on the display. Note, for example, that the size of the basic texture element in the right-hand figure appears



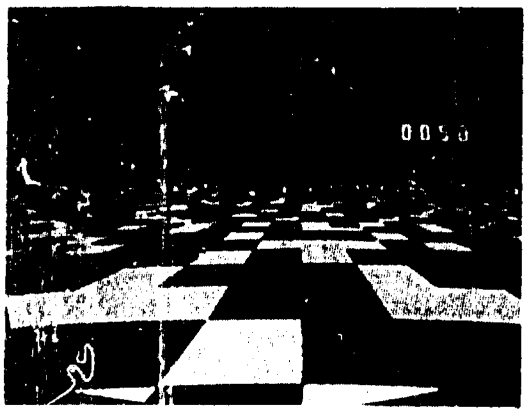
(a) Four-foot cell size.



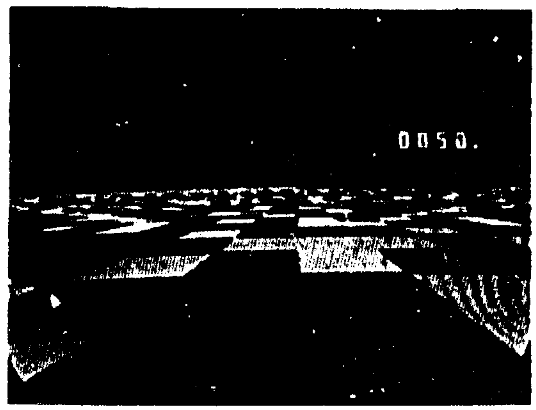
(b) Eight-foot cell size.



(c) Sixteen-foot cell size.



(d) Thirty-two-foot cell size.



(e) Sixty-four-foot cell size.

Figure 7. Illustration of the Effect of Increasing Cell Size While Holding All Other Parameters Constant. Starting with a cell size of four feet, cell size was progressively increased by a factor of two.

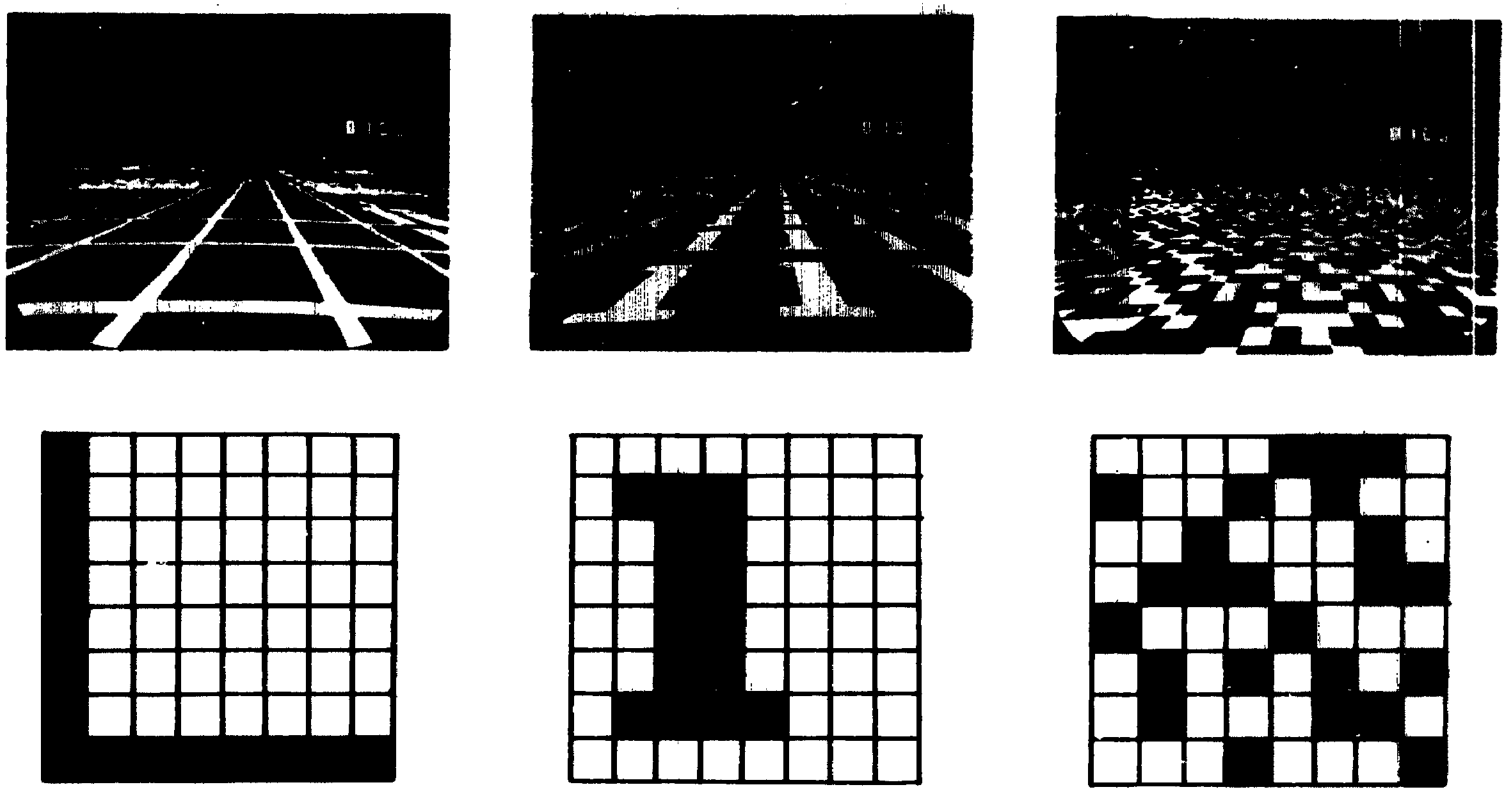


Figure 8. The Line Drawings Illustrate the Construction of Three Types of Ground Texture Elements. The photo located above each drawing shows the appearance of the ground plane when each of these three types of texture are used.

clearly smaller than the elements in the remaining two patterns. This is true despite the fact that a greater number of cells were used to construct the former.

Several other features associated with texture patterns are illustrated in figure 8. The right-hand pattern (which was constructed to approximate a random pattern) is more dense and less symmetrical appearing than the other two patterns. The location of perspective lines (formed by the texture elements) is less clear in the right-hand pattern. The qualities of symmetry, density, and clearness of perspective lines appear to be important attributes of ground texture and warrant experimental investigation in the future.

<u>Levels of Texture and Texture Breakpoints</u>: Calculation will show that texture elements small enough to provide useful cues at low altitudes appear so small as to be indistinguishable at very high altitudes. And, conversely, texture elements large enough to be seen at high altitudes would subtend such a large visual angle at low altitudes that the angular extent of a single element would exceed the display field of view--with the effect that the ground plane would appear as an untextured surface. The same problem exists for slant range. That is, small elements cannot be seen at great distances and large elements, whose size is suitable for great distances, exceed the display field-of-view when the viewing distance is small. Thus, providing appropriate texture cues over a wide range of altitudes and slant ranges is a problem that required special consideration in the design of the VCAD.

The difficulty was overcome by constructing the ground texture from a hierarchy of four nested texture patterns. The basic texture element for each step of the hierarchy has a unique shape and, at each succeeding step up the hierarchy, the size of the texture element increases by a factor of eight. The first order texture element is constructed from an 8x8 matrix of cells. The second order texture element is constucted from a 64x64 matrix of cells or, stated differently, from an 8x8 matrix of first order matrices. Similarly, the third order texture element is constructed from an 8x8 matrix of second order matrices. Thus, a fourth order texture element is 512 times as large as a similarly shaped first order element. The drawings in figure 9 are examples of unique texture shapes that can be utilized for the four orders of texture. Furthermore, these drawings illustrate the relative size of the elements of each order of texture.

Additional understanding of the hierarchical structure of texture elements and the concept of "breakpoints" can be obtained by referencing the series of photos shown in figure 10. These five photos illustrate changes in the display as a result of increasing altitude from 10 to 1000 feet--where pitch angle and roll angle are held at zero and cell size is maintained at 8 feet. Note that only the first order texture elements ("1"s) can be seen at an altitude of 10 feet. As altitude is increased to 50

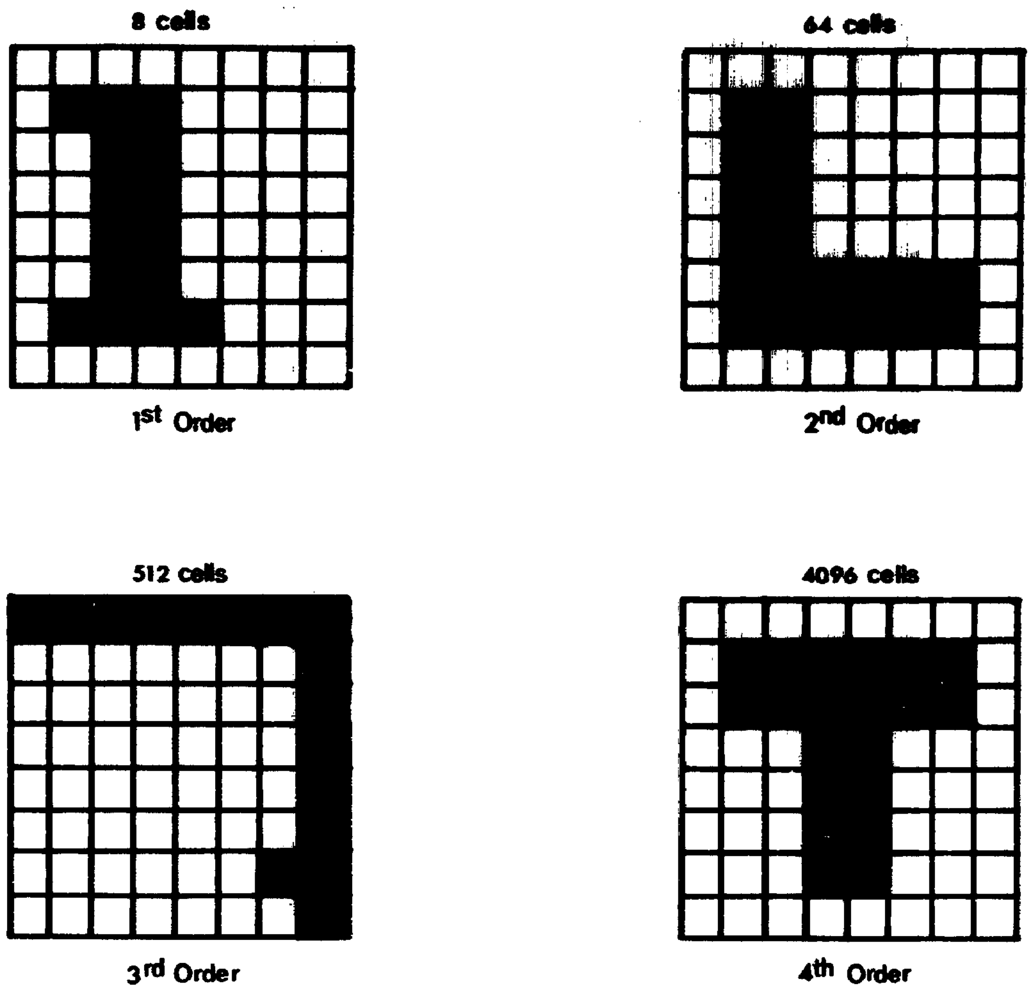


Figure 9. Illustration of the Construction and Relative Size of the Four Orders of Texture Elements.

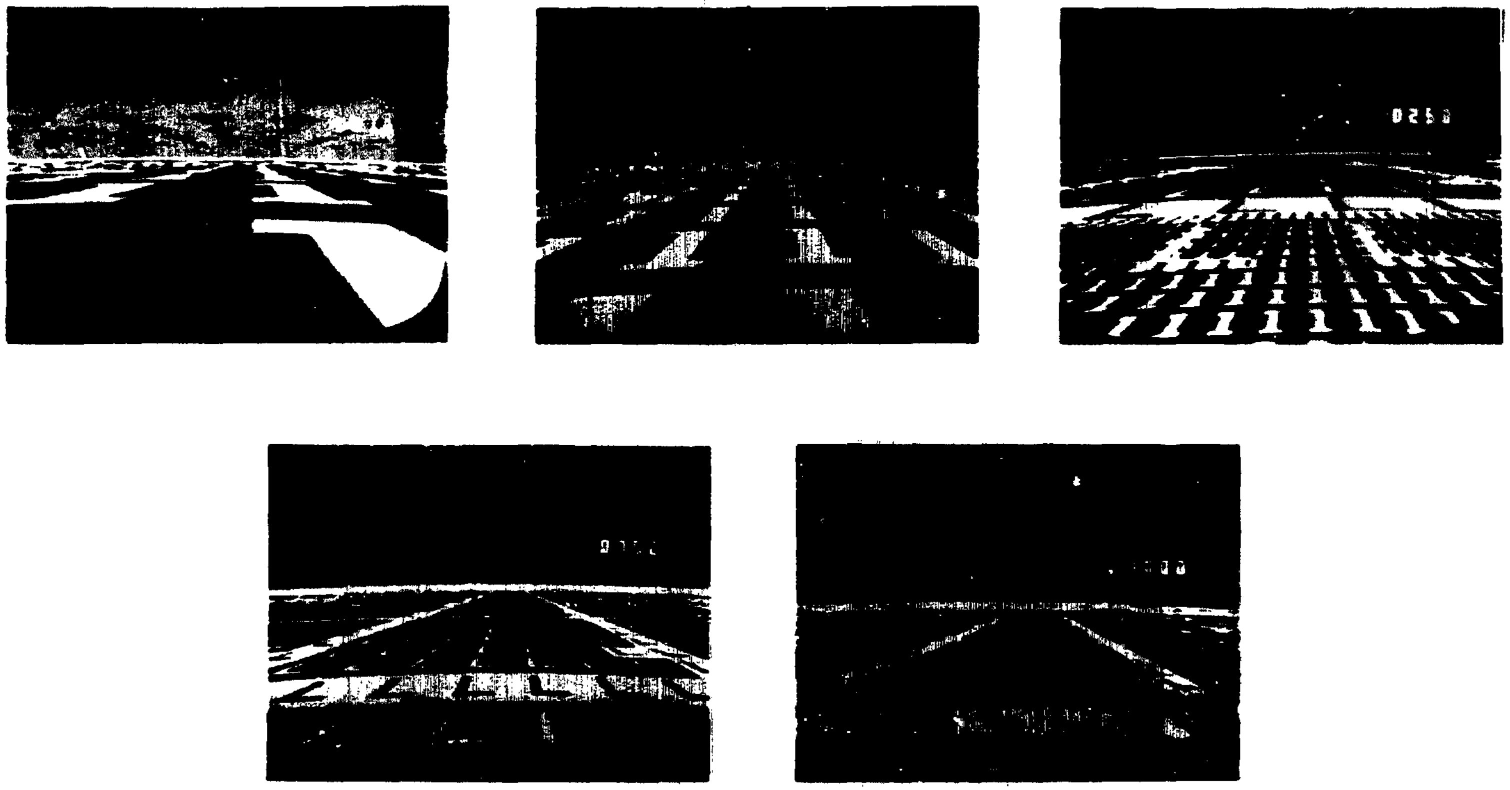


Figure 10. Illustration of the Emergence of Texture Breakpoints as a Function of Increasing Altitude.

feet several changes occur: the angular size of the "1"s decreases, the second order texture ("L"s) begins to emerge, and a point can be seen in the distance at which the first order texture abruptly terminates. This abrupt line of demarcation between orders of texture is what is referred to as a breakpoint. Breakpoints always occur at a constant slant range (for a given cell size and display viewing angle) regardless of the orientation of the aircraft. The designers of the CADG computed the slant range at which the texture elements of a given order could no longer be resolved by the display and programmed the CADG to disregard elements in this order of texture when the slant range to these elements exceeded the preassigned value. This design feature is comparable to the acuity threshold of the human eye, except that the CADG is not susceptable to the variety of conditions which affect the eye's acuity threshold. Only second order texture elements can be seen beyond the first breakpoint, though the second order texture elements are embedded in the first order texture. Actually, all of the higher orders of texture are embedded in this manner, although this fact is not apparent from an altitude of 50 feet.

As altitude continues to increase, the first breakpoint systematically approaches the bottom edge of the screen until at 1000 feet this breakpoint can no longer be seen. This means that the slant range to the nearest ground point appearing on the display exceeds the slant range to the first breakpoint. If the aircraft were pitched down and if the altitude did not exceed the slant range to the first breakpoint, the first breakpoint would reappear on the display.

At 1000-feet altitude, the second breakpoint and the third order texture can be clearly seen. If altitude were increased still further, the fourth order texture would emerge and the second and third breakpoints would disappear in the same manner as the first breakpoint.

Display Viewing Angle: Display viewing angle has been discussed previously and was defined graphically in figure 2 (see angle α). It is display viewing angle that determines the angular extent of the model world that is projected on the display and is defined quite independently of the eye of the observer. It should again be emphasized that one-to-one correspondence with the "real world" exists only when the visual-angle subtended by the display is the same as the display viewing angle.

Zenith Marker: A unique 8x8 cell matrix is located on the sky plane at the intersection of the nadir-zenith line and the sky plane. This matrix permits the introduction of a uniquely shaped symbol at this point on the sky plane making this geographical point easily distinguished. This symbol, hereafter referred to as the Zenith Marker, is always located directly above the simulated aircraft regardless of its orientation and provides a valuable cue to pitch when positive pitch angle is so large that the horizon line does not fall within the display field of view.

Special Symbology: A variety of information supplements in the form of symbols can be added to the basic display as desired. As only one of these symbols were utilized in the present experiments, they deserve only brief mention here. The special symbols that can presently be presented on the display include: (a) a flight-path which can be presented in either an earth-stabilized or flight-director mode, (b) three-dimensional building-shaped objects located on the ground plane, (c) a symbol in the shape of a cross which identifies the position of the impact point, (d) square-shaped symbols whose position on the display can be made to be a function of one or more of the flight parameters, (e) a fixed-scale moving-pointer cursor which can be made to display error in any of the flight parameters, (f) two sets of four-digit numerics which can be made to reflect the momentary value of any two flight parameters, (g) a uniquely textured ground location which maintains a fixed geographical position on the ground plane, and (h) a runway whose length and directional orientation can be varied.

CADG Inputs: The inputs accepted by the CADG include: north rate, east rate, altitude rate, heading angle, pitch angle, and roll angle. It should be noted, however, that roll is not computed by the CADG but is accomplished by rolling the yoke on the neck of the CRT.

Fixed-Base Cockpit

Experimental subjects were housed in a general purpose fixed-base cockpit that was designed to accommodate a variety of different display/control configurations, depending upon the requirements of the experiment. An inside and outside view of the fixed-base cockpit is shown in figure 11. The display used in all the experiments discussed in this report was the "basic VCAD," which includes the textured sky and ground planes, but none of the special symbols previously discussed. Manipulation of display size was accomplished through changing the size tube on which the display was presented while holding viewing distance constant at 32 inches. Four different tube sizes were used. Although these are referred to in this report as 5-, 8-, 14-, and 17-inch displays (commercial designation) their actual height and width dimensions are as shown below. Shown in parenthesis is the vertical and lateral visual angle subtended by each of the four tubes. Unless viewing distance is the same, only the visual angle dimensions should be generalized to other situations.

Tube Size	Height	Width
5"	3.4" (6°6')	4.5" (7° 58')
8"	5.4" (9°34')	7.0" (12° 30')
14"	9.1" (16° 16')	11.9" (21° 4')
17"	10.9" (19° 22')	14.2" (25° 2')

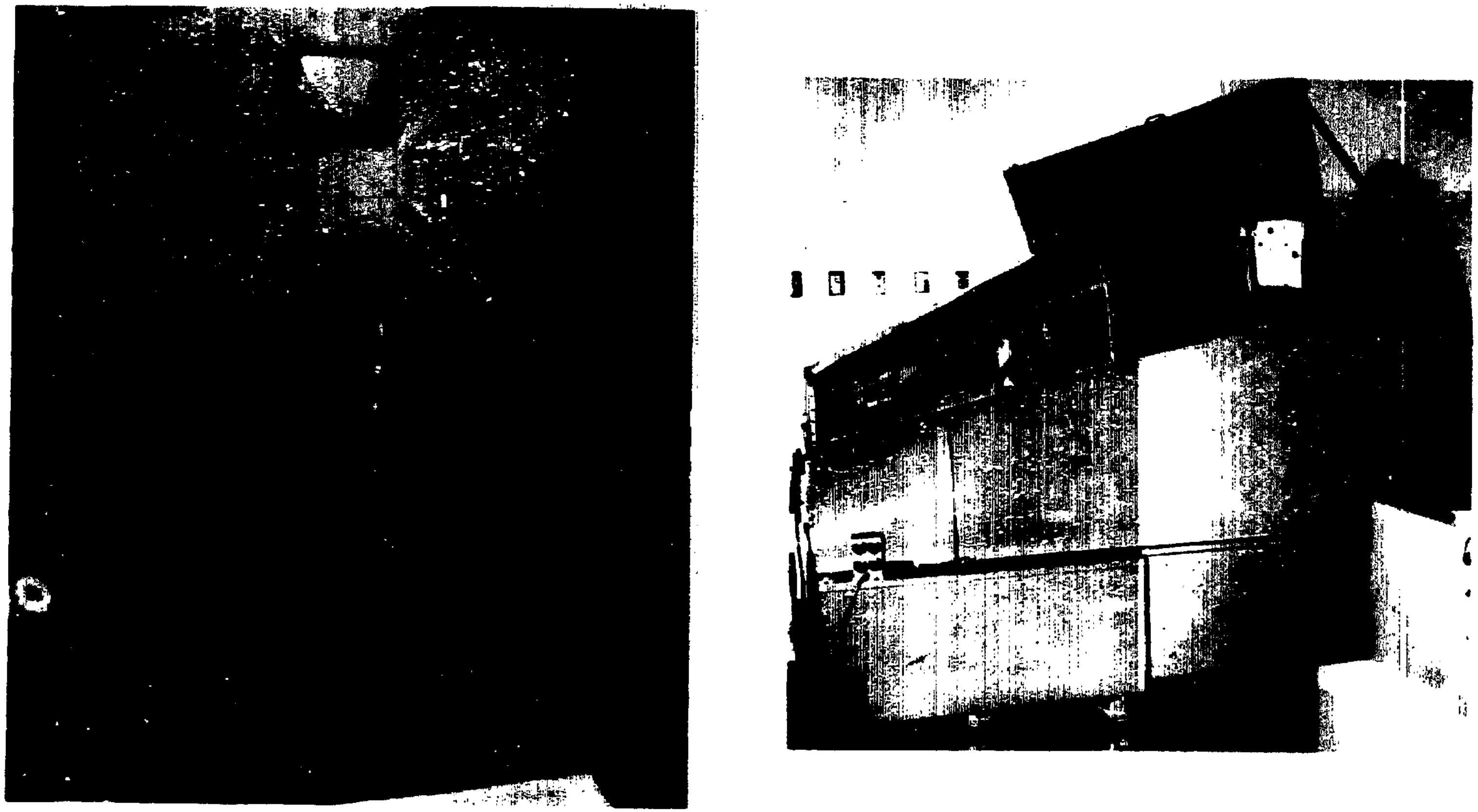


Figure 11. Inside and Outside Views of the JANAIR Fixed-Base Cockpit.

A standard aircraft joystick located a convenient distance in front of the subject and in alignment with the mid-line of his body, was the control device used in all the experiments. Through manipulation of the joystick, the subject controlled the output of a potentiometer which, in turn, drove the value of the flight parameter being investigated. The joystick served as a rate control in all experiments; i.e., the rate of change of the controlled parameter being a function of the amount of displacement of the joystick from its null position.

Control of roll angle was accomplished through lateral displacement of the joystick. The display/control relationship used was the same as that in an aircraft; for example, displacement of the joystick to the right of the null position caused the horizon to roll toward the left. Longitudinal displacement of the joystick controlled either altitude or pitch, depending upon which parameter was being investigated. When pitch judgments were being investigated, displacement of the stick forward of the null position caused the aircraft to pitch down, whereas, displacement in the opposite direction caused the aircraft to pitch up. In those experiments dealing with altitude judgments, a forward stick displacement caused altitude to decrease and a backward displacement caused altitude to increase. It should be understood that only one parameter was under the S̲s control during any given experiment.

Performance Measurement System (PMS)

The PMS is composed of (a) an Electronic Associates, Inc. (EAI) TR-48 analog computer used in conjunction with an EAI general purpose Digital Expansion System (DES-30), (b) a 24-channel Minneapolis-Honeywell Visicorder, and (c) a Cimron Model 703 Data System. The components of the PMS are functionally connected with both the CADC and the fixed-base cockpit so that the value of each of the flight parameters as well as stick output can be continuously monitored by the analog computer. The raw quantities can be directly read-out on the digital voltmeter located on the TR-48 computer, they can be fed to the Visicorder and read-out in the form of continuous oscillographic records, and (or) they can be fed to the Cimron Data System which can sequentially scan, evaluate, and print-out the value of up to 100 inputs. Typically, however, the signals were processed in some way before being read-out. For example, error scores were computed by subtracting the actual value of a parameter from an assigned command value (hereafter referred to as the "standard"). The majority of the performance measures were derived by submitting the error output to various mathematical manipulations. In the case of discrete judgments, such as setting altitude at some prescribed value, the PMS provided a discrete read-out of error, and various computations were performed on off-line computers. In the case of continuous judgment tasks, such as maintaining one of the parameters at a constant value over an extended period, error was read-out continuously and the computations performed on-line by the TR-48 analog computer. The various performance measures obtained are discussed below for both the discrete and the continuous case.

Performance Measures

Average Absolute Error (AAE)

Average Absolute Error for the continuous and the discrete case is given in equations 7 and 8, respectively.

$$\frac{1}{t}\int_{t1}^{t2} |e(t)|dt \qquad (7)$$

$$\frac{1}{n}\sum_{i=1}^{n} |ei| \qquad (8)$$

These measures are particularly useful in comparing relative performance under various experimental conditions. However, because the same amount of error can be accumulated in a variety of different ways, assuming total equivalency of performance on the basis of equivalency of AAE scores alone may result in important qualitative differences in performance being overlooked.

Root Mean Square (RMS) Error

RMS error can be shown to be the standard deviation of the error about the designated standard. If the error is normally distributed about the standard, that is, if no response bias exists, one can expect error to be within one standard deviation of the standard 68 percent of the time. The equations for RMS error for the continuous and the discrete case are given below in equations 9 and 10.

$$\sqrt{\frac{1}{t}\int_{t1}^{t2} e^2 dt} \qquad (9)$$

$$\sqrt{\frac{1}{n}\sum_{i=1}^{n} e^2 i} \qquad (10)$$

Average Error (AE)

Equations 11 and 12 show, for both the continuous and discrete case, that AE is obtained by accumulating error with respect to the sign and represents the mean of the error distribution. Thus, AE provides a direct index of performance bias.

$$\frac{1}{t} \int_{t1}^{t2} e(t)dt \tag{11}$$

$$\frac{1}{n} \sum_{i=e}^{n} ei \tag{12}$$

Standard Deviation about Mean of Error Distribution (SD_E)

Whereas RMS is an index of the variability of error about the assigned or "true" standard, SD_E can be thought of as the variability of error about S's concept of memory of the standard--which may or may not correspond with the true standard. If S's concept of the standard does correspond with the true standard and error is distributed normally, AE (bias) will be zero and RMS and SD_E will be identical. If S's performance is biased, RMS would necessarily increase in size yet SD_E could remain small. Such a result would indicate a consistent bias in S's concept of the standard, but a small amount of variability about this "assumed" standard. A large RMS score coupled with a similarly large SD_E score would simply indicate highly erratic performance. Thus, SD_E serves primarily as a diagnostic tool.

$$\sqrt{(RMS)^2 - (AE)^2} \tag{13}$$

Time History Recordings

Continuous oscillographic records of the momentary value of the flight parameter being investigated and momentary error were obtained. Only selected trials were recorded. Because quantitative measures obtained from oscillographic records are typically inaccurate and extremely time consuming to obtain, the primary purpose of these recordings was to provide a qualitative notion of response strategies and to provide an additional check of the validity of the analytical measures described above.

Analytic Approach

Customarily, the analytic approach is discussed in the "Results" section of a research report. In the present case, however, the analytic approach for all the experiments is essentially the same; and devoting a separate section to their discussion avoids unnecessary repetition. Also, because the analytic approach that was developed is one which will be unfamiliar to many readers, it is believed necessary to both describe the approach in general and to discuss some of the conditions which necessitated its development.

Experimental Designs

The experimental designs found most suitable for present purposes were Fractional Replications designs using repeated measures. The principle advantage of these designs is that large savings in time are realized by not having to obtain data for all possible combinations of conditions--as is required when using Complete Factorial designs. Fractional Replications designs have the disadvantage of not providing estimates of certain third- and higher-order interactions. In most cases these higher-order interactions were judged to be highly unlikely or negligible in size. When this assumption could not be made with a high degree of confidence, Complete Factorial designs were used.

The use of repeated measures permit the separation of subject effects from treatment effects thus providing a more precise measure of treatment effects. Also, Repeated Measures designs permit the estimation of treatment effects despite the presence of sequence effects such as learning and fatigue--both of which were present in these experiments.

Polynominal Response Surface Model

A central assumption of Analysis of Variance models, which have traditionally been used for Fractional Replications designs with repeated measures, is that the shape of the performance curve is the same for all subjects and all conditions (Winer, 1962). This assumption is seldom valid for psycholcgical data (see Thorndike, 1949, p 113) and was certainly invalid for the data reported here. An even more serious shortcoming of conventional models is that they do not permit generalization beyond the particular conditions investigated. Either of these shortcomings would, in itself, render Analysis of Variance models unacceptable for present purposes.

A class of models called "Response Surface Models" were found to be compatible with our research objectives and yet did not require that unrealistic assumptions be made about the form of data. These models serve to define continuous relationships between values of the independent and dependent variables. Thus, the value (or "response") of the dependent variable can be calculated for any combination of values for the independent variables. A "surface" can be generated from a systematic collection of "responses"--hence, the name response surface.

The most fruitful type of Response Surface Models have used polynomial series to approximate functions and are called "Polynomial Response Surface Models" (Cochran & Cox, 1957). These models rely on the fact that many differently shaped functions can be closely approximated by polynomial series. Moreover, the models can be fitted through the use of well established regression techniques. To illustrate the form of response surface models, an example of a second order (quadratic) surface involving the

independent variables X_{1i} and X_{2i} and the dependent variable Yi is given below:

$$Yi = \beta_0 + \beta_1 X_1^2 + \beta_2 X_1 + \beta_3 X_1 X_2 + \beta_4 X_2 + \beta_5 X_2^2 + e_i \tag{14}$$

Similar equations can be written for extensions of the form involving several variables and/or high order (powered) terms.

Analytic Procedure

The first step in the analytic procedure was to generate a population of prospective predictor variables; the most obvious candidate being the independent variable of primary interest--in this case, display size. Moreover, any other independent variable that was experimentally controlled and which is likely to contribute to judgment error must be included as a prospective predictor variable, whether or not the variable is of particular interest in the study. Examples of such variables include the skill of the individual S ("subject") the stage in learning or amount of fatigue ("sessions") the value of the parameter being judged ("standard"), and so on. Also, powered terms such as $(\text{sessions})^2$ or $(\text{sessions})^3$, and interaction terms such as "sessions x subjects" are prospective predictors and must therefore be included. The population of prospective predictor variables thus derived averaged about thirty for the various experiments.

This population of predictor variables was then submitted to a Computerized Stepwise Multiple Regressions Analysis (developed by Wherry, 1966) which sequentially evaluates each predictor variable. The technique selects terms for a regression equation with the requirement that each term selected must improve the "expected fit" of the model to the data.[5] The end result is a print-out of those variables found to contribute to the prediction of the dependent variable and their associated weighting constants, the composite of which is a multi-variate regression equation. This equation is a highly versatile tool and provides the capability of expressing the relationship between the dependent variable and independent variables singly or in any desired combination.

The equations generated for the various experiments were in turn, used to generate predicted values of the dependent variable for a number of values of the independent variable over the range of interest. When plotted in graphical form, these data clearly illustrate the dependent/independent variable relationship, if such a relationship were indicated by the raw data. It should be noted, however, that in only one case was the data extrapolated beyond the range of independent variables actually investigated.

- - - - - - - - - - - - -

[5] The "expected fit" criteria differs from the more familiar "significance test criteria" usually used with multiple-regression techniques. Its intent is to maximize the correlation of the predicted data with data obtained in identical subsequent experiments. Conversely, the "significance test" criteria is concerned with the likelihood that the variance associated with a particular term is due to experimental error.

It should also be noted that all the graphs represent predictions for the average S. Those graphs showing the relationship between the dependent variable and the display size represent predicitions for a point late in the experiment where S's condition mastery was at its maximum.

A standardized format was developed for summarizing the results of the multiple-regression analyses for the various experiments. An example of this format is shown in figure 12 and is explained below. Each row of the matrix shown in figure 12 represents the results of the analysis for the particular dependent variable identified in the left-hand margin. In all cases, more than one independent variable was analyzed.

The classes of predictor variables selected are identified in the upper margin of the columns to the left of the double line. The number of individual variables falling in each class is shown in the associated cell of the matrix. To illustrate the meaning of "classes" of predictor variables, if display size (display size)^2 and (display size)^3 were all selected as predictor variables, all three will logically fall into a "class" which could reasonably be termed "display size." Thus, display size would be shown adjacent to one of the columns and the number "3" shown in the corresponding cell below. The reader is cautioned against attaching special meaning to the order of presentation of the variable classes or to the number of variables within each class as neither has any bearing on the relative "importance" of the predictor variables. The relative importance of the variable classes and the individual variables within these classes can be determined only by examining plots of the predicted data directly.

The total number of predictor variables included in the regression equation is shown in the first column to the right of the double line in figure 12. The second column to the right of this line shows the total number of data points (experimental trials) upon which the analysis was based.

Proceeding to the right, the next two columns list an F-ratio and show the probability with which an F-ratio of this size would have occurred by chance. The F-ratio was obtained by dividing the ratio of the "variance accounted for" over its degrees of freedom by the ratio of the "residual" over its degrees of freedom. Its significance indicates the unlikeliness of the results due to chance fluctuations.

The column labeled $\overline{R}^2$ is interpreted as the proportion of variance in the dependent variable accounted for by the composite of predictor variables in the regression equation. $\overline{R}^2$ is obtained by subtracting the ratio of the Residual and Total sums of squares from 1, and its interpretation is analogous to the interpretation of error SS in analysis-of-variance designs. In the broadest sense, $\overline{R}^2$ is an index of the degree to which the model fits the data. In order to provide an index for evaluating the

	IV_1 . . . . IV_n				Total Predictors	Total Data Points	F-ratio	p	$\bar{R}^2$	$\bar{R}$
DV_1										
DV_2										

Figure 12. Example of the Format Used to Summarize the Results of Multiple-Regression Analyses.

magnitude of $\bar{R}^2$, one should consider that the "average" experiment reported in the psychological literature accounts for roughly 10 percent of the total variance.

The multiple correlation coefficient ($\bar{R}$) is shown in the right-hand column of figure 12. $\bar{R}$ is obtained by taking the square-root of $\bar{R^2}$ and is an index of the validity of the regression equation.

ROLL ANGLE EXPERIMENTS

The overall objective of the set of three experiments reported in this section was to: (a) define the relationship between display size and error in roll angle judgments and (b) to obtain reliable estimates of the absolute accuracy with which roll angle can be judged from the basic VCAD. The task in the first two experiments was, essentially, a tracking task which required Ss to detect and null roll angle error, where error is defined as the momentary difference between the roll angle appearing on the display and an assigned "standard" roll angle. A forcing function was used in both experiments which continuously forced roll angle off the standard. All other flight parameters remained constant in the first experiment. It was discovered, however, that this experimental arrangement provided unrealistically powerful cues to roll angle--cues that would not be available in an actual flight situation. For this reason, the experiment was replicated with a forcing function also being introduced into pitch angle, although S still maintained control of only roll angle. This procedural change forced S to adopt cues that are generalizable to a complex flight situation.

The third experiment in this set measured the accuracy with which Ss could set roll angle to each of a number of values. Whereas the first experiments were concerned with S's ability to continuously detect and null error from a standard, the third experiment was concerned with S's ability to make absolute judgments of the magnitude of roll angle.

Roll Angle Maintenance Experiments

Subjects

Four laboratory personnel were used as Ss All Ss were initially flight naive, having had no prior flight experience nor any previous experience judging the value of flight parameters from the VCAD. The same four Ss were used in both experiments so their level of learning was substantially higher during the second experiment.

Apparatus

These experiments were conducted in the laboratory facility described in the General Apparatus section of this report. The special features specific to these experiments are described below.

1. The joystick in the fixed-based cockpit was modified to drive roll angle only. Movement of the stick to the right of the null position caused the horizon line to tilt toward the left, whereas, movement to the left of the null position caused the horizon line to tilt toward the right. This display-control relationship is directly analogous to that in an aircraft. The rate of change in roll angle was a function of the amount of displacement from the null position.

2. A forcing function was introduced into the roll angle channel in both the original and the replicated experiment. A forcing function was also introduced into the pitch angle channel during the replicated experiment although S had no control of pitch. The magnitude and frequency of the sine wave forcing function driving roll angle was 2.6 cpm and ± 15 degrees respectively.[6] Roll angle could be made to oscillate about a reference point which was systematically varied throughout the experiment. The sine wave forcing function driving pitch caused pitch angle to oscillate about zero pitch at a rate of 2.0 cpm and a magnitude of ±15 degrees.

3. A 5-inch display was not available at this time so only 8-, 14-, and 17-inch displays were investigated.

4. The performance indexes obtained included AAE, RMS error, AE and oscillographic records of momentary error and momentary roll angle on the first and the last two trials of each sessions.

Procedure

The S's task in these experiments was to maintain a standard roll angle (either 0, 20, or 60 degrees) over a two-minute trial using either an 8-, 14-, or 17-inch display. Each S performed ten such trials during each of nine sessions. The size of the display and the roll angle standard remained constant during any one session, but varied

- - - - - - - - - - - - - - - -

6 Low frequency and low amplitude forcing functions were deliberately selected to ensure that the limiting feature of the perceptual-motor task was the perceptual portion rather than the motor portion of the task. This procedure ensures that any difference in task proficiency is a function of S's ability to decode the display rather than his ability to encode his response. It was originally feared that use of an entirely periodic forcing function might allow the subject to memorize spacial-temporal aspects of the forcing function and thus be able to null error without referencing the display. This possibility was checked by asking highly trained Ss to perform the task for one minute in the ordinary manner and for the next minute with their eyes closed. Although Ss were able to reproduce the amplitude with a fair degree of accuracy, their timing very quickly became out of phase with the forcing function with the result that error far exceeded that obtained on any trial during the experiment.

systematically from session to session. The order with which each of the nine possible combinations of display size and roll angle standard were administered is shown in table 2. This order constitutes a Fractional Latin Square design and was obtained by randomly selecting four rows from a standard 9x9 Randomized Latin Square. This design ensures that each display size is represented at least once during each of the nine sessions and further ensures that each S receives each combination once during the experiment.

Prior to the first trial the nature of the task was carefully explained to the Ss and they were given the opportunity to become familiar with the display/control relationship. Prior to each session roll angle was set on the appropriate standard. The Ss were told that the roll angle which presently appeared on the display was the standard which they were to maintain. They were further instructed to spend the inter-trial interval trying to memorize the standard. Subjects were given a verbal ready signal and verbal signals to indicate initiation and termination of each trial. During the inter-trial interval, which was approximately one-minute in length, the experimenter recorded the error data from a digital voltmeter.

Results

The results of the first experiment showed that after only a moderate amount of practice, Ss were able to maintain all of the three roll angle standards with the same remarkably high degree of accuracy regardless of the size of the display being used. Peak error seldom exceeded two degrees during a trial and the absolute error averaged over a trial was less than one-half degree. These findings were immediately suspect and led to a careful examination of the strategies used by the Ss It was discovered that, without exception, Ss sought out some type of extraneous reference point on either the face of the display or on its edge which happened to be in alignment with the horizon line, a breakpoint line, or a perspective line when roll angle was in the standard position. On subsequent trials they simply proceeded to maintain alignment of these two arbitrary reference points.

Because a different external reference point was required for each display size and each roll angle standard, a wide variety of extraneous items were used as reference points. These included the corners of the display, minute smudges and specks on the tube face, small scratches and smudges on the tube mounting, etc.

In considering the legitimacy of this strategy it became apparent that it would be appropriate only when pitch angle remained constant--a condition that would rarely exist in a flight situation for any significant period of time. For this reason the study was judged invalid for its

Table 2. Order of Presentation of Display Size and Standards in Roll Angle Maintenance Experiment

Subject		Session (Order) 1	2	3	4	5	6	7	8	9
S_1	Display Size*	14	17	8	14	8	14	8	17	17
	Standard **	0	60	20	60	60	20	0	20	0
S_2	Display Size	8	8	17	17	14	8	17	14	14
	Standard	20	0	60	0	0	60	20	60	20
S_3	Display Size	14	14	17	8	17	8	8	14	17
	Standard	60	20	0	20	20	0	60	0	60
S_4	Display Size	17	17	14	14	8	17	14	8	8
	Standard	20	0	0	60	0	60	20	60	20

* Display size in inches.
** Standard in degrees.

intended purpose, and the decision was made to replicate the experiment using a forcing function in pitch angle as well as roll angle. This procedure forced S to adopt strategies that would be appropriate in an actual flight situation.

Although invalid for its intended purpose, one finding of the study is worthy of mention. The tendency of Ss to search out external references and their reluctance to rely upon the internal cue of "memory" of the standard appears to be universal. Whether this tendency is due to the inherent features of the visual mechanism or due to past experience with conventional displays remains to be answered.

The results of the multiple regression analysis for the replicated roll angle experiment are summarized in figure 13. It can be seen that analyses were performed for three different criteria, viz., AAE, RMS, and AE. It was found that AAE and RMS error correlated very highly (.97) indicating that the same trends were manifest in both measures. It was also found that AE was effectively zero by the senventh session, indicating that there were no consistent biases and that errors were symmetrically distributed about the standard. Given these two findings, RMS error can be confidently assumed to be a reliable estimate of the standard deviation of the error about the standard. Since standard deviation units are the most easily translatable, predicted data for only RMS error is presented in the following graphs.

First note that the multiple regression equation that evolved from the analysis is composed of six classes of predictor variables and a total of 11 individual predictor variables. The highly significant F-ration (33.6 with 11 and 359 df) indicates that this weighted combination of predictor variables accounts for a highly significant proportion of the total variance. The proportion of variance accounted for ($\bar{R}^2$ = 51 percent) and the multiple correlation coefficient ($\bar{R}$ = .71) indicate that the equation is an effective and valid estimator of RMS error. Every indication suggests that the predicted scores derived from the regression equation are highly generalizable.

Examining the classes of predictor variables shown in figure 13, it can be seen that display size-- the variable of primary interest--was selected as a predictor variable. This finding indicates that judgment error is indeed dependent upon the size of the display being used. The selection of the standard x display size interaction indicates that the effect of display size on judgment error differs, depending upon the magnitude of the roll-angle standard. The nature and magnitude of these effects are illustrated in figure 14, which shows RMS in minutes of arc as a function of display size and standard roll angle. These curves represent predictions to a point late in learning (session) for the average S. The curves for the 20-inch and 5-inch displays represent extrapolations

	Display Size	Standard	Sessions	Subjects	Standard x Display Size	Standard x Sessions	Subject x Sessions	Total Predictors	Total Data Points	F-ratio	p	$\bar{R}^2$	$\bar{R}$
AAE	1	2	1	2	1	1	3	11	360	47.2	$\leq .005$	.59	.77
RMS	1	2	2	2	1	1	2	11	360	33.6	$\leq .005$	.51	.71
AE	1	3	2	3	1	1	4	15	360	18.1	$\leq .005$	.43	.66

Figure 13. Summary of Multiple-Regression Analysis for Roll Angle Maintenance Experiment Number Two.

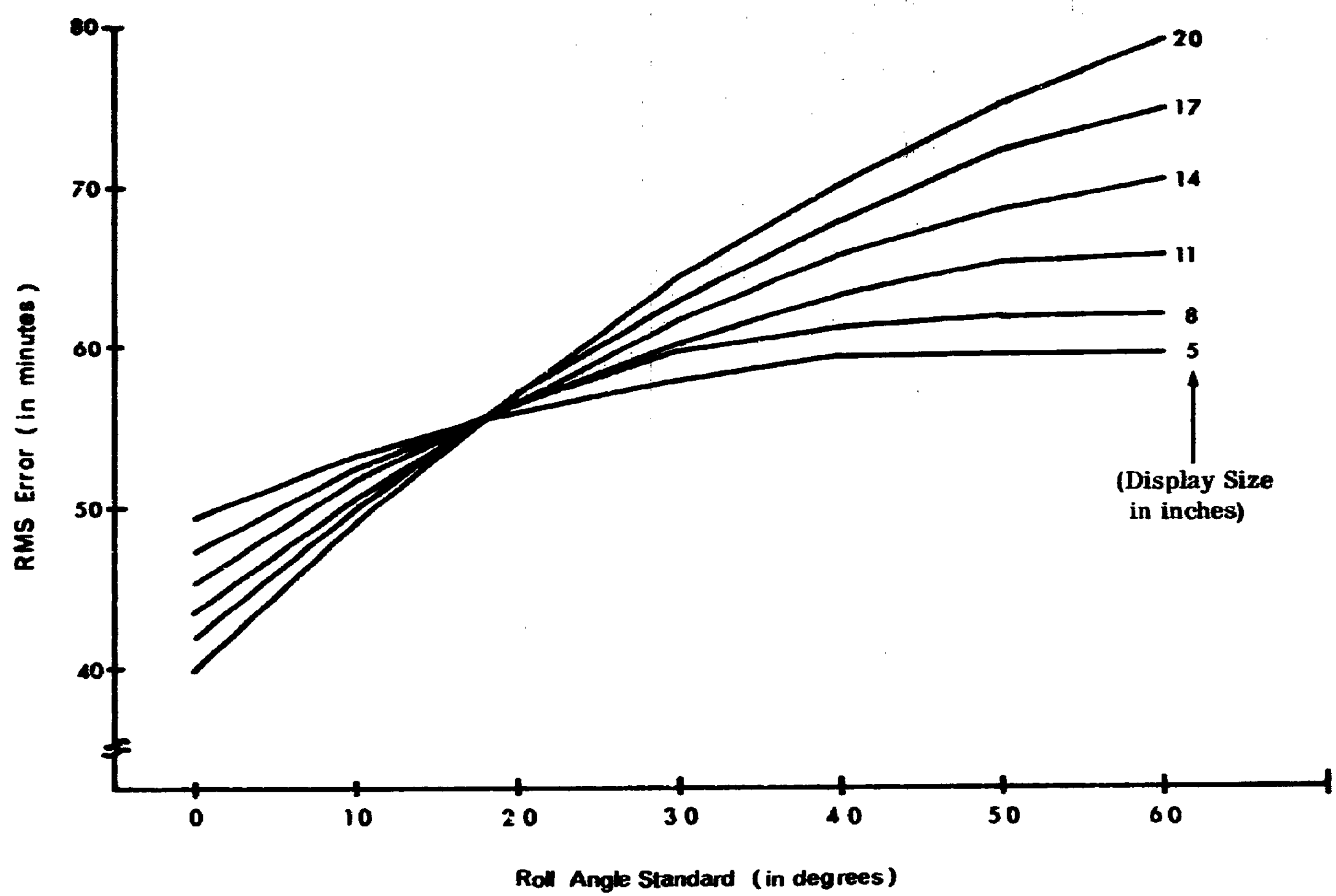


Figure 14. Accuracy of Roll Angle Maintenance as a Function of Display Size and Roll Angle Standard. Predicted data for the average S on session 7.

beyond the range of display sizes investigated so they should be interpreted with reservation. A high degree of confidence can be placed on all other predicted points however.

Figure 14 shows that the "optimal" display size is dependent upon the magnitude of the roll angle being judged. If the roll angle standard falls below about 18 degrees, larger size displays are favored; if the roll angle standard exceeds 18 degrees, smaller sizes prove optimal. Considering the data overall, an 11-inch display would appear to be the most effective trade-off for minimizing judgment error over a wide range of roll angles.

The magnitude of the difference between the best and the worst condition in figure 14 can be seen to be relatively small. Error is found to be minimum when a 20-inch display is used to judge a roll angle standard of zero degrees and is maximum when the same display size is used to judge 60-degree roll angles. Although the difference between these two conditions (about 40 minutes of arc) is highly statistically significant, it would appear to have little practical importance.

Turning to the question of the absolute accuracy with which one can expect roll angle to be maintained, it is necessary to establish a level of confidence. If one defines an interval as 3 standard deviation units on either side of the roll angle standard, it can be assumed that roll angle can be maintained within this interval 99.74 percent of the time. Applying this confidence criterion to the data for the 11-inch display, one can state that at a 60-degree roll angle, the average S should maintain roll angle within ± 3.15 degrees of the standard 99.74 percent of the time and that he should never perform worse than this regardless of the roll angle he desires to maintain.

Although not of primary interest, several of the other predictor variables should be discussed. A predictor variable for Ss indicates that individual Ss performed differently and a subject x sessions predictor variable indicates that different Ss learn at different rates. This effect is the reason why traditional analysis of variance models were inappropriate for these data.

Selection of a sessions predictor variable indicates that judgment error changed as a function of practice. The sessions x standards predictor variable indicates that this change was not the same for all roll angle standards. These effects are graphically illustrated in figure 15 which shows RMS error as a function of sessions and roll angle standard. These curves represent predicted data for the 11-inch display and the average S. The slight upturn of the two lower curves is an artifact resulting from fitting a quadratic equation to the data. It can safely be assumed that the error score for session 7 represents the true asyptotic level of these conditions. The important trend in figure 15 is that the

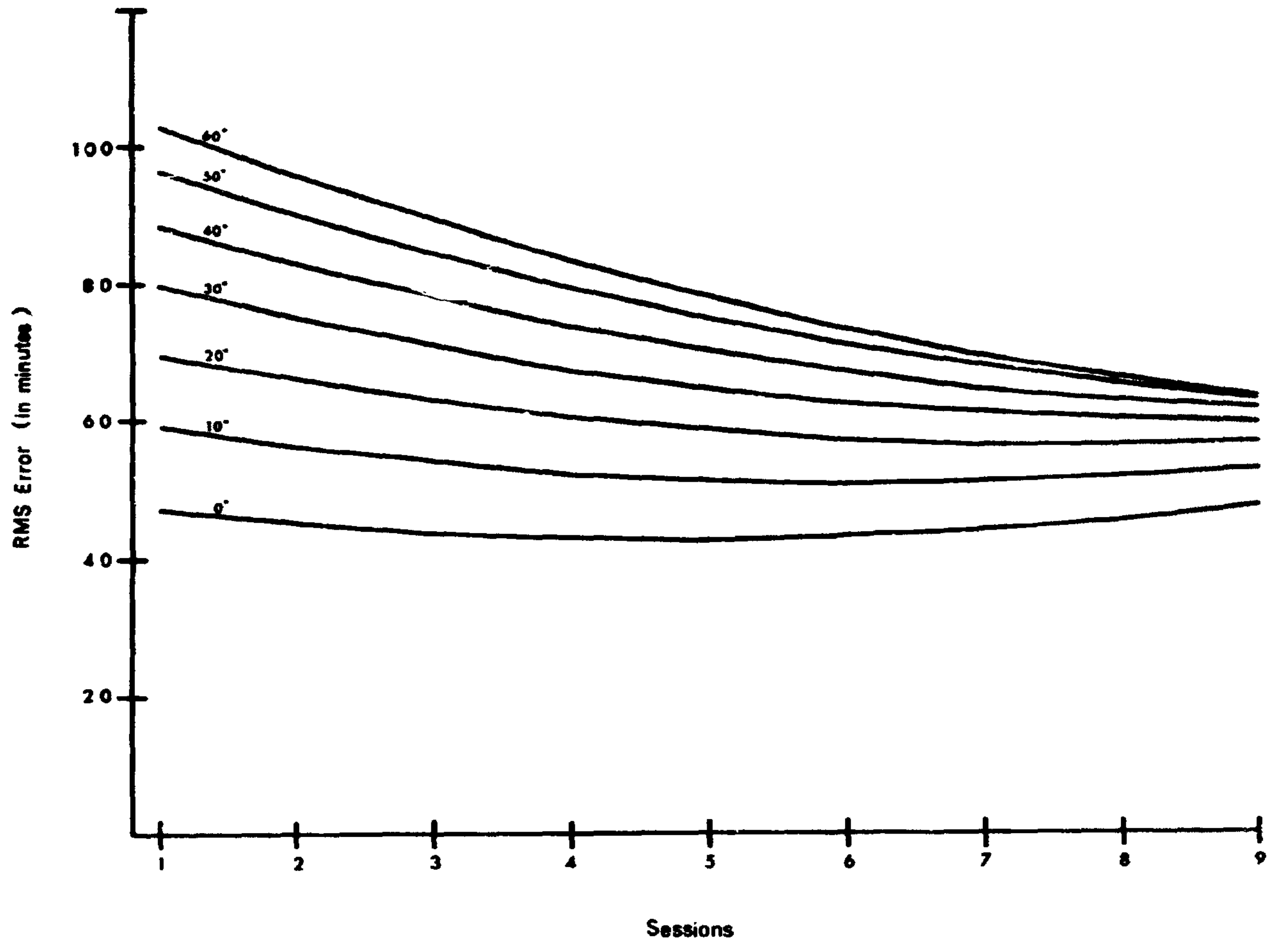


Figure 15. Accuracy of Roll Angle Maintenance as a Function of Sessions and Roll Angle Standard. Predicted data for the 11-inch display and the average <u>S</u>.

curves tend to converge with practice. Although all Ss were highly practiced by the ninth session. It is possible that the differences between roll angle standards would have eventually "washed-out" with sufficient training.

Discussion

The prediction that the accuracy of maintaining a fixed roll angle would increase as a simple function of display size was not confirmed by these data. No satisfactory explanation for the curious interaction between display size and roll angle standard is available at this time, although a number of possible contributory effects can be hypothesized. The difference between roll angle standards is likely to be a function of S's past experience. He has had more experience judging the horizontalness of objects, so it seems reasonable that zero degrees would be the easiest condition regardless of display size. Since zero degrees is familiar to S, and since the tip of the horizon line moves through greater excursions for larger displays, it also seems reasonable that larger displays would be favored at small roll angles.

To explain why smaller displays are favored at large roll angles, one must consider that (a) although proficiency in the task is a function of S's ability to detect and null temporary changes from the standard, some long term drift in error is bound to occur and (b) the ability to null this long term drift is a function of S's ability to learn and recall the standard. Accepting these facts, it seems reasonable to hypothesize that because larger displays subtend such a large visual angle, a clear image of the composite configuration would be more difficult to attain.

The surprising degree of accuracy with which roll angle could be maintained suggests that the basic VCAD may be more self sufficient than had previously been supposed.

Absolute Judgment of Roll Angle

The key objective of this experiment was to determine the accuracy with which absolute roll angle judgments can be made with the basic VCAD, and to determine whether judgment accuracy varies as a function of display size. In the previous experiments task proficiency was a function of S's ability to detect change in roll angle whether or not he could attach any numerical value to the change. The present experiment is concerned with measuring S's ability to extract quantitative information regarding roll angle from the basic display.

This experiment also served a secondary objective. The question was raised as to whether the use of non-pilots in these basic experiments would result in valid data. Although the roll angle maintenance task was deliberately designed to minimize the effects of past experience and to

minimize complex skill requirements, the possibility existed that experienced pilots who had had considerable experience judging roll angles in a flight situation might have developed more skill in making absolute roll angle judgments than could be developed by non-pilots, even after considerable practice. Flight experience was therefore introduced as an additional independent variable in this study.

Subjects

Three experienced Naval aviators and three non-pilot laboratory personnel served as S's in this experiment. The pilot population had no previous experience with the VCAD. The non-pilot S's had served in the previously described experiment concerned with roll angle maintenance. However, they had no familiarity with ascribing numerical values to roll angle other than the three roll angle standards used in the roll angle maintenance study, viz., 0-, 20-, and 60 degrees.

Apparatus

The apparatus used in this experiment was essentially the same as that utilized in the roll angle maintenance experiment except that the forcing functions was not introduced into the roll angle channel as before. Functionally, the apparatus gave S the capability of adjusting roll angle to any desired magnitude by manipulating the control stick in the lateral plane. Roll angle could be maintained at a stable position by simply returning the stick to its null position after roll angle had reached the desired position.

The PMS was modified to provide a direct read-out of the roll angle on the digital voltmeter. The voltmeter presented roll angle rounded to the nearest degree.

Procedure

The method by which absolute roll angle judgments were measured in this experiment is as follows: Twenty-five angles, varying from 60 degrees left roll to 60 degrees right roll in 5-degree increments were selected as roll angle standards. The experimenter sequentially presented these standards to S by verbally calling out the direction and the value of the standard; e.g., "15 degrees left roll," "25 degrees right roll," etc. The S then proceeded to adjust roll angle to what he judged to be the prescribed standard, signaling the experimenter as to when he had completed his adjustment by calling "mark." The experimenter then recorded S's actual roll angle setting and proceeded to the next standard. Although this was a self-paced task, S's were instructed to "not expend an inordinate amount of time in making a judgment."

A trial in this experiment is defined as one roll angle judgment; and a block of trials consists of 26 judgments. Each of the 25 roll

angle standards was judged once during a block of trials except for the zero-degree standard which was judged twice. The order of presentation of the standards was arranged randomly within blocks.

The initial session, hereafter designated as session zero, consisted of three blocks of trials in which S's were given no information as to the accuracy of their judgments. This "no feedback" condition was necessary in order to obtain accurate estimates of the relative skill of pilots and non-pilots prior to learning taking place. Thereafter, non-pilot S's performed six additional sessions of two blocks each; receiving feedback as to their actual setting immediately after each trial. Although the difference in pilot's and non-pilot's judgment accuracy prior to learning was the primary concern, the pilots consented to participate in 3 additional blocks of feedback trials. Feedback was given by verbally informing S's of the magnitude of their setting in degrees.

Roll angle judgments were measured under three conditions of display size; viz., 8-, 14-, and 17-inch displays. The order of presentation of display size is illustrated in table 3. The row adjacent to each S number lists the sequential order with which that S received the three displays. This design ensures that each S receives the displays with equal frequency and ensure that each display size is used equally often at each stage of practice.

Results

The results of the multiple regression analysis are shown in figure 16. It will be recalled that AAE is the average of the absolute difference between the actual and the standard setting, whereas AE is the average of this difference when summed with respect to sign. The value of the standard was subtracted from that of the actual settting so that a positive difference indicates that the setting was too large and a negative difference indicates that the setting was too small.

Figure 16 indicates that the regression equations for AAE and AE, both composed of 13 predictor variables, accounts for a statistically significant portion of the total variance (F-ratios of 30.5 and 13.7 for AAE and AE respectively). The regression equation derived for AAE is seen to account for 19.5 percent of the total variance and that for AE, 9.8 percent. These values indicate the presence of considerably more experimental error that was found in the previous experiment. Upon subsequent examination of the experimental procedures it seemed likely that a large portion of this error resulted from measuring judgment error to only the nearest degree accuracy. Although this fact does not invalidate the results of the experiment, it may tend to produce slightly conservative estimates of judgment accuracy. That is, judgments may be made with slightly greater accuracy than is indicated by the regression equation.

Table 3. Order of Presentation of Display Sizes in Absolute Roll Angle Judgment Experiment

	No Feedback			Feedback											
Session	0			1		2		3		4		5		6	
Block	1	2	3	1	2	3	4	5	6	7	8	9	10	11	12
Subject															
S_1	14*	17	8	14	14	14	14	17	17	17	17	8	8	8	8
S_2	17	8	14	17	17	17	17	8	8	8	8	14	14	14	14
S_3	8	14	17	8	8	8	8	14	14	14	14	17	17	17	17
S_4	14	17	8	14	17	8									
S_5	17	8	14	17	8	14									
S_6	8	14	17	8	14	17									

* Display size in inches.

	Flight Experience	Standard	Sessions	Subjects	Standard x Sessions	Subject x Sessions	Standard x Experience	Total Predictors	Total Data Points	F-ratio	p	$\bar{R}^2$	$\bar{R}$
AAE	1	4	2	2	0	2	2	13	1638	30.5	$\leq .005$	.195	.44
AE	1	3	0	1	2	4	2	13	1638	13.7	$\leq .005$	.098	.31

Figure 16. Results of Multiple-Regression Analyses for Absolute Roll Angle Judgments.

Considering the sets of predictor variables chosen by the analysis, the most significant finding is that the display size variable did not prove to be a predictor of either AAE or AE. On the basis of this finding, it can be concluded that display size had no measurable effect on the accuracy with which absolute roll angle judgments can be made.

Another independent variable of interest was "Flight Experience" (pilots vs non-pilots). It can be seen that flight experience proved to be a predictor of both AAE and AE. The effects of both flight experience and sessions (practice) on AAE is illustrated in figure 17. The curves in figure 17 represent predicted data for the 30-degree standard which, as will be seen later, is the standard at which AAE was maximum. These curves are therefore to be considered the "worst case" condition. It was surprising to find that judgment error for pilots is considerably greater than that for non-pilots. Evidently, the knowledge of the position of 0-, 20-, and 60 degree standards obtained from having served as S's in the roll angle maintenance experiment was sufficient to give non-pilots this edge. Since there is no reason to believe that non-pilots would be inately superior at this type of judgment task, it would seem certain that error for the two groups would have converged to the same level had practice been continued. It is also worthy of note that the learning curve for non-pilots appears to have nearly leveled off by session 6, although some slight improvement would still be expected with additional practice.

The effects of roll angle standard upon AAE is illustrated in figure 18. This figure shows AAE as a function of the magnitude of the roll angle standard for both pilots and non-pilots. Note, however, that the curve for pilot performance is predicted for session 1 while the curve for the non-pilot performance is predicted for session 6. Since no difference was noted for roll direction, these curves can be considered representative for both left and right roll angles.

The similarity in shape of the pilot and non-pilot curves, despite the large difference in the amount of practice under feedback conditions, suggests that the two curves would probably be the same given equivalent practice. The finding that AAE is very small at zero degrees and increases as a function of the magnitude of the roll angle standard is not particularly surprising. That is, judgment of horizontalness and verticalness is a common requirement in everyday experience, but one is seldom faced with having to quantify the amount of slope or tilt of objects without the aid of a measuring device.

The fact that judgment error differs so greatly from standard to standard makes it difficult to make a simple quantitative statement regarding the accuracy with which roll angle can be decoded from the basic display. To facilitate the formulation of such statements, a second scale is provided on the right-hand margin of figure 18 making it possible to read error for each of the standards in standard deviation units (note

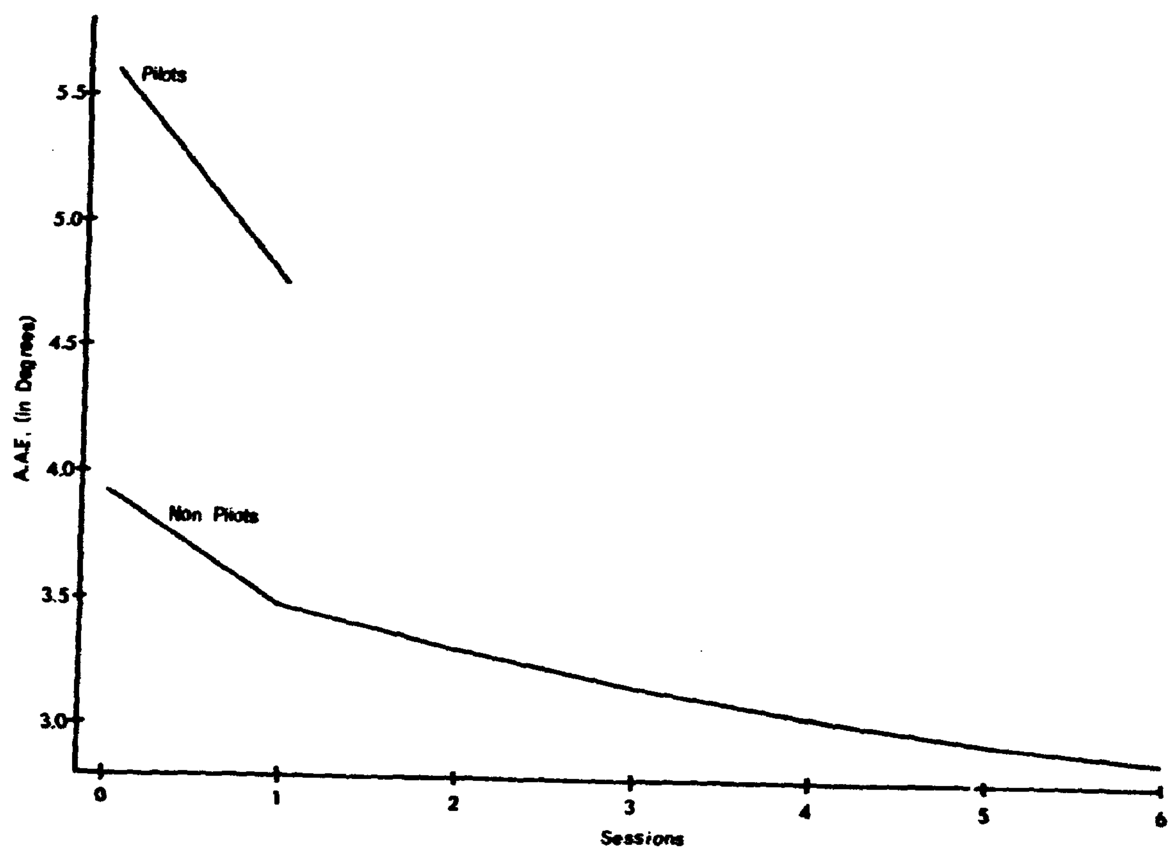


Figure 17. Average Absolute Error in Roll Angle Judgments Across Sessions. Predicted data for 30-degree standard and for pilot and non-pilot Ss.

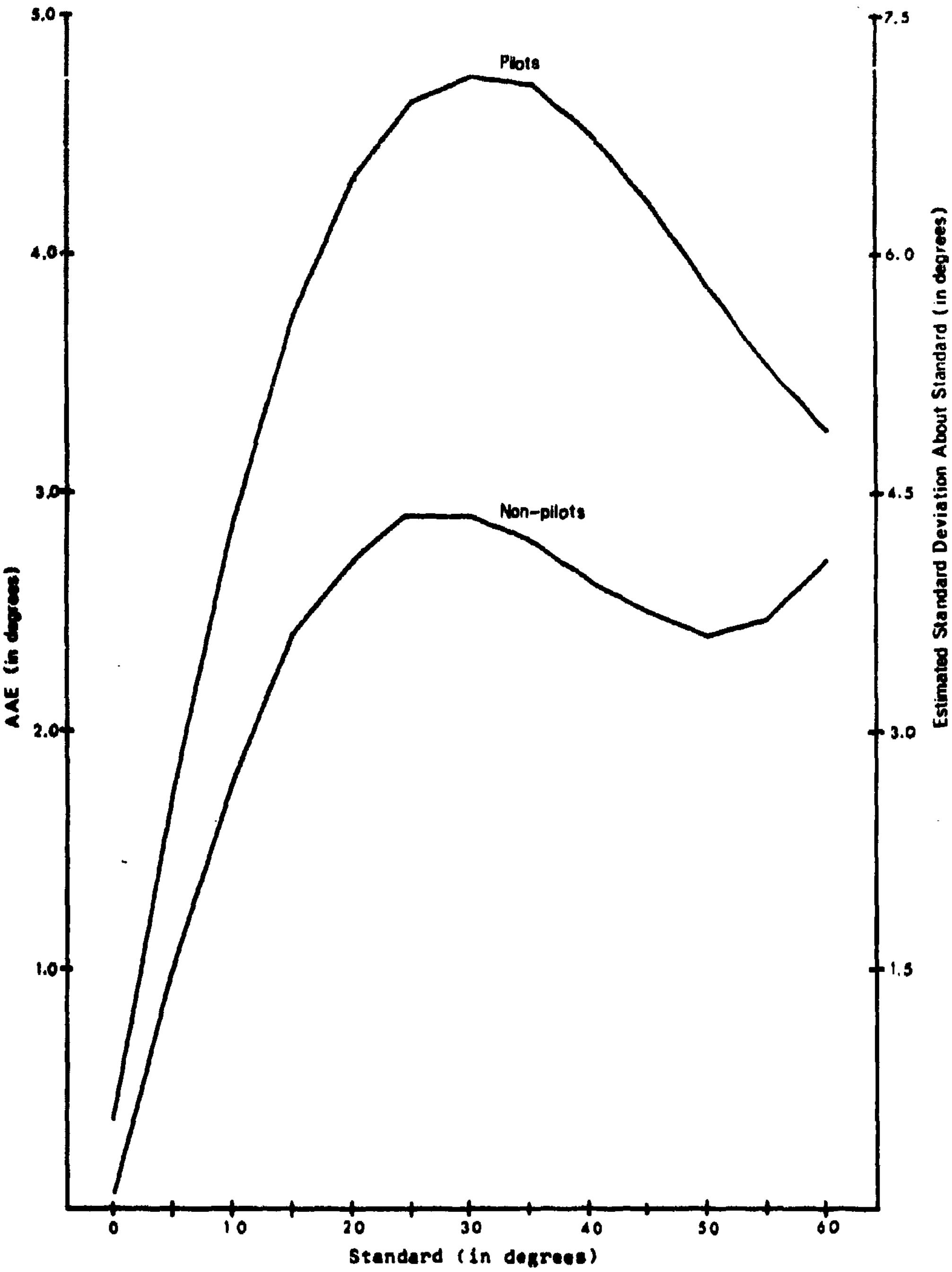


Figure 18. Average Absolute Error of Roll Angle Judgments as a Function of Roll Angle Standard. The pilot data are predicted for session 1 and the non-pilot data are predicted for session 6.

the equality: SD = 1.5 (AAE)). As an example of the manner in which figure 18 can be used in making statements regarding absolute decoding accuracy, assume one is interested in the accuracy with which the 20-degree standard can be expected to be judged. Figure 18 shows the standard deviation for the 20-degree condition to be approximately 3.9 degrees. Assuming that errors are normally distributed about the standard and assuming one wishes to attach a high degree of confidence to his statement, say the 99 percent level, then the statement can be made that the "average" S will adjust roll angle to within ±3 (3.9) = ±11.7 degrees of the standard in 99.97 percent of his attempts.

Figure 19 illustrates the relationship between AE--a measure of judgment bias--and magnitude of the roll angle standard. The curve in figure 19 represents predicted AE for non-pilots after six training sessions. Except for the zero-degree standard, a positive score indicates that the average setting was larger than the corresponding standard. Conversely, a negative score indicates that the average setting was smaller than the standard. For the zero-degree standard, a positive AE score indicates a right roll, and a negative score indicates a left roll. For example, the positive .3-degree AE score for the zero-degree standard indicates that when S's judge roll angle to be zero, the horizon line is actually tilted to the left an average of .3-degrees.

Figure 19 shows that a small amount of judgment bias exists but that the bias is not consistent across all roll angle. Angles less than 45-degrees tend to be overestimated. That is, the average S tends to perceive his settings as being larger than they are in reality. Angles larger than 45 degrees tend to be underestimated, suggesting that Ss perceive large angles to be smaller than they are in reality. Maximum negative bias (.7 degrees) is seen to occur at the 25-degree standard, and maximum positive bias (.95 degrees) at the 60-degree standard. Although the presence of this bias invalidates the assumption that errors are normally distributed about the standard, the bias is so small for most conditions that the estimates of decoding accuracy derived from figure 18 will be affected by only a small amount.

A second prediction made earlier was derived from the fact that the magnitude of the excursion through which the "tip" of the horizon line travels as a function of an increment in roll angle is greatest at a roll angle of 39 degrees. It was predicted that accuracy should tend to improve as the roll angle standard approached 39 degrees. Although a dip in the error curve was noted in the vicinity of 39 degrees, the dip appeared to be symmetrical about a minimum of 50 degrees.

Discussion

The results of this experiment confirmed the prediction made earlier that display size--within the range of interest here--would have no

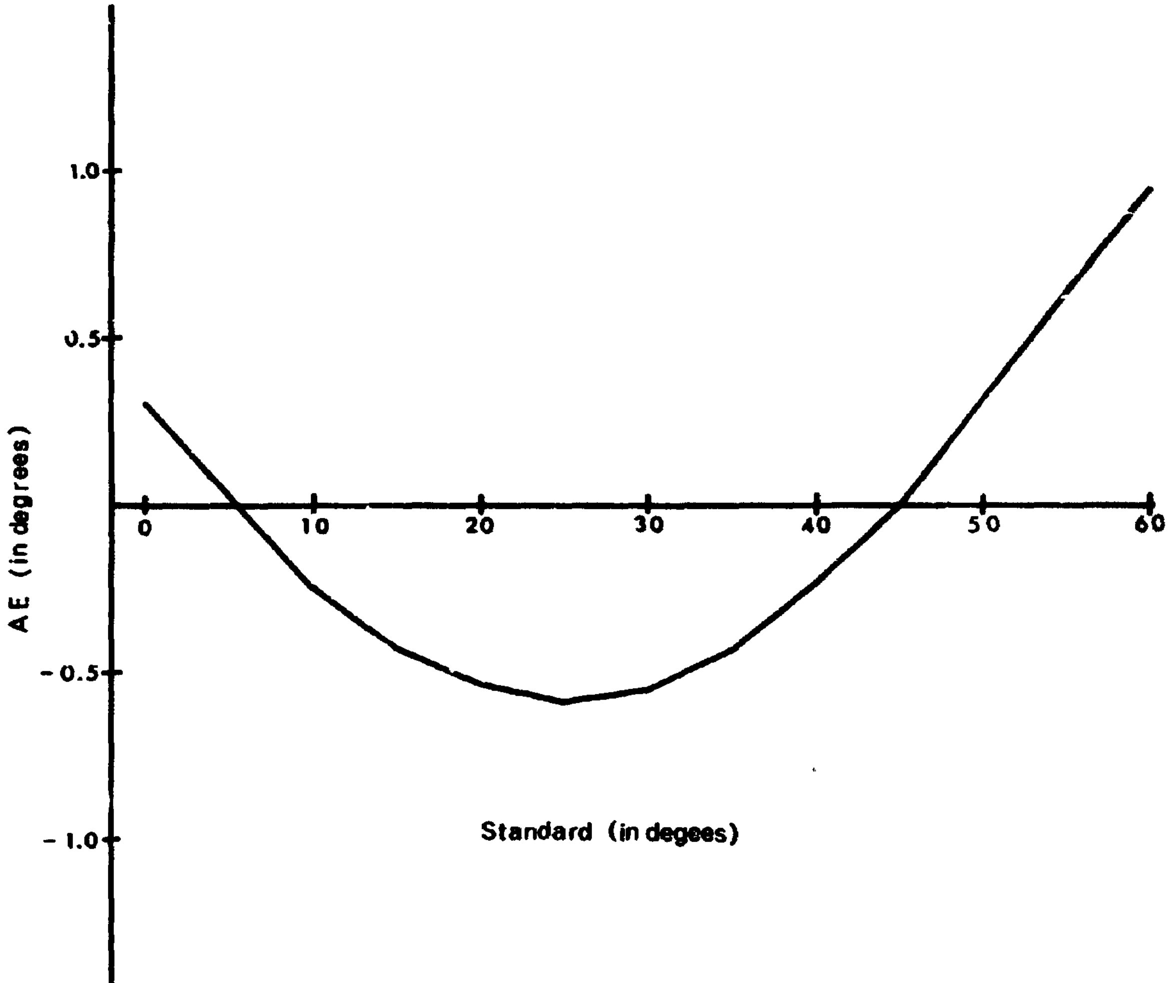


Figure 19. Average Error as a Function of Roll Angle Standard. Predicted data for non-pilot population on session 6.

effect on the accuracy with which absolute roll angle judgments can be made. A finding that was not predicted was the superior initial performance of non-pilots over pilots. It was stated previously that this difference was undoubtedly due to the non-pilots becoming familiarized with the 0-, 20-, and 60-degree standards in a previous experiment. Evidently knowledge of these three standards enabled these Ss to interpolate to intermediate standards with a fair degree of accuracy. It is nevertheless surprising that pilots had not established similar internal references to roll angle magnitude during their flight experiences. This may suggest that little learning of this type will take place so long as external references such as the reference marks on ADI displays are immediately available.

Although average bias was small, it is interesting to speculate as to why the bias curve was of its particular shape. One possible explanation is as follows: In making a setting, S must reference roll angle magnitude to some point with which he is familiar. The most likely candidates would appear to be zero or 90 degrees. If S references zero degree and his natrual tendency is to overestimate the difference between his setting and his reference point, then his setting will tend to be too small. Similarly if he references 90 degrees and displays the same tendency to overestimate the difference between his setting and his reference, then his settings will be too large. If S referenced zero-degree when judging angles less than 45 degrees and referenced 90 degrees when judging angles larger than 45 degrees one would expect precisely the bias curve presented in figure 19.

ALTITUDE EXPERIMENTS

The purpose of the two experiments reported here was to define the nature of the relationship between display size and the accuracy of altitude judgments, and to obtain reliable estimates of the absolute accuracy with which altitude can be judged from the VCAD. Two types of judgment skill were investigated. The first experiment was designed to measure the type of judgments utilized in maintaining altitude at a constant level and is referred to as the altitude maintenance experiment. Proficiency in this task is dependent upon S's ability to detect and to null change in altitude regardless of whether he is able to assign numerical values to this change. The second experiment was designed to measure S's ability to derive quantitative information regarding altitude from the basic VCAD. This experiment is termed the absolute judgment experiment.

Since the powerfulness of altitude cues varies a great deal as a function of altitude, it was necessary to measure altitude judgments over some specified range of altitudes in order to insure generalizability of

results. Since altitude judgment accuracy is most critical at low altitude, it was obviously necessary to measure judgment accuracy at low altitude levels. The decision as to the upper limit of the altitude range was based upon a mathematical analysis of altitude cues which has been reported previously (Cross, 1968). This analysis was concerned with the amount of change in ground texture size (expressed in terms of the visual angle subtended by a ground texture element) and in the apparent movement of ground texture elements (expressed in terms of angular velocity) that results from a given increment in altitude. The results of this analysis showed that once altitude reaches 1000 feet, the size and the apparent velocity of ground texture elements change very little with further increases in altitude. That is, these cues to altitude are very insensitive at altitudes higher than 1000 feet. For these reasons the range of altitude selected for investigation was from 10 to 1000 feet.

Altitude Maintenance Experiment

Subjects

Four male laboratory personnel served as Ss in this experiment. All of these individuals had participated in at least one previous VCAD experiment, but none had had any prior experience in judging altitude. All Ss were generally familiar with the manner in which the display changes as a function of altitude change.

Apparatus

The reader is referred to the "General Apparatus" section of this report for a discussion of the apparatus used in this experiment. The following items represent features not previously discussed which were unique to this experiment.

1. The output of the joystick was set-up to drive altitude rate in the CADG. Movement of the stick to a position forward of the null position introduced a negative rate, causing altitude to decrease. A stick position in back of the null position (toward S) introduced a positive rate, causing altitude to increase. The rate of change of altitude was a function of the magnitude of the stick displacement.

2. The display sizes investigated in this experiment included a 5-, 8-, 14-, and 17-inch display.

3. During the experimental trials, altitude was continuously driven off a standard altitude by a modulated sine wave forcing function. The average frequency of this forcing function was 2.0 cpm and its amplitude was ±150 feet (from the standard altitude).

4. Pitch angle, roll angle, and heading angle remained at zero degree throughout the experiment. Longitudinal ground speed was set at 150 knots with no lateral movement (drift) being introduced.

5. The equipment was modified to provide the experimenter with the capability of adjusting altitude to any one of the five altitude standards that were investigated. These five standards were 10, 50, 100, 500, or 1000 feet. The forcing function caused altitude to vary symmetrically about whichever of the standard altitudes that had been set into the equipment. Altitude automatically returned to the standard at the termination of each trial.

6. The performance measures obtained included integrated error (IE), integrated absolute error (IAE), mean square error, and time-history recordings of momentary stick position, momentary altitude, and momentary altitude error.

Procedure

After having been seated in the fixed-base cockpit, Ss were instructed that their task during the experiment was to maintain altitude at a constant level by appropriately manipulating the joystick in the longitudinal axis. After being informed of the nature of the display/control relationship, Ss were invited to manipulate the stick in order to get a "feel" for its effect on altitude. All Ss were thoroughly familiar with the VCAD so it was unnecessary to explain the changes in the display that result from altitude change.

Prior to the first trial, altitude was set to the appropriate standard, and the S instructed that "what he saw on his display was the first altitude he would be attempting to maintain." He was permitted to examine the display configuration until he reported feeling certain that he could recall the appearance of that standard. He was also instructed that altitude would automatically return to the standard at the termination of each trial and that he should spend the time between trials trying to better learn the appearance of the display when at the standard altitude. Subjects were given unlimited time to study the display each time a new standard was introduced.

Performance was measured on each of five altitude standards (10, 50, 100, 500, and 1000 feet) paired with each of four size displays (5, 8, 14, and 17 inch) giving a total of 20 experimental conditions. All 5 altitude standards were investigated during each daily session. The order with which these standards were investigated during each daily session was established by means of a randomization procedure. Subjects were given 6 consecutive two-minute trials with each of the five altitude standards--giving a total of 30 trials per day. In order to provide a better estimate of learning effects, the 6 trials were divided into two blocks of 3 trials each for computational purposes. All trials within a daily session were performed on the same size display.

The order of presentation of the four displays was dictated by two augmented Latin Squares. Table 4 illustrates the order with which each

Table 4. Order of Presentation of Display Sizes During Altitude Maintenance Experiment

	Sessions															
	1		2		3		4		5		6		7		8	
Blocks	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
Subjects																
S_1	17*	17	14	14	8	8	5	5	17	17	14	14	8	8	5	5
S_2	14	14	8	8	5	5	17	17	14	14	8	8	5	5	17	17
S_3	8	8	5	5	17	17	14	14	8	8	5	5	17	17	14	14
S_4	5	5	17	17	14	14	8	8	5	5	17	17	14	14	8	8

* Display size in inches.

S received each of the four displays. Each S performed the experimental task with each of the four displays on two temporally separate occasions. Moreover, this design insures that each display size is represented equally often at each stage in practice. The entire experiment entailed 240 two-minute trials per S, giving a total of 960 experimental observations.

Results

The RMS and IE scores were used to compute the standard deviation of altitude error about the mean of the error distribution (SD_E). It was found that SD_E and RMS were essentially equivalent measures, their correlation being .985. As was mentioned previously this finding indicates the absence of judgment bias and indicates that RMS error may be considered a reliable estimator of the standard deviation of an error distribution whose mean is zero. That is, the mean of the altitude distribution is equivalent to the altitude standard. For this reason, the results of the multiple regression analyses are presented for AAE and RMS error only. The results of these analyses are summarized in figure 20.

Figure 20 shows that the regression equation derived for AAE and RMS error are composed of 14 and 23 predictor variables, respectively. Both regression equations were found to account for a highly significant proportion of the total variance ($p < .005$). Examination of the $\overline{R}^2$ values will show that both equations were found to account for an unusually high proportion of the total variance--78 percent and 75 percent for AAE and RMS, respectively. These summary data indicate that a high degree of confidence can be placed upon the predictions derived from the regression equations.

The two independent variables of primary interest in this study are display size and altitude standard. Figure 20 shows that both were selected as predictors of both the performance measures. This finding is evidence that both variables influence the accuracy of altitude judgments as measured by the altitude maintenance task. Another predictor variable of interest is the Display Size Standard interaction variable. Selection of this variable indicates that the effect of display size on judgment accuracy depends in part upon the altitude standard being judged. The remaining predictor variables, although of academic interest, have little bearing upon the main objective of this experiment.

Figure 21 shows predicted AAE (in volts) for the average S on session 7. Predicted AAE is shown as a function of both display size and magnitude of the altitude standard. (The break in the ordinate of figure 21 should be noted.) It is clear from figure 21 that AAE generally increases as a direct function of magnitude of the altitude standard. Plotting AAE against altitude standard on Semi-Logarithmic paper showed that this relationship was only roughly logarithmic in nature.

	Display Size	Standard	Sessions	Subjects	Display Size x Standard	Standard x Sessions	Subjects x Sessions	Total Predictors	Total Data Points	F Ratio	P	R^2	R
AAE	2	4	2	2	1	3	0	14	960	248.6	$\leq .005$	.78	.88
RMS	2	4	4	2	1	6	4	23	960	127.1	$\leq .005$	.75	.87

Figure 20. Summary of Multiple Regression Analysis for Altitude Maintenance Experiment.

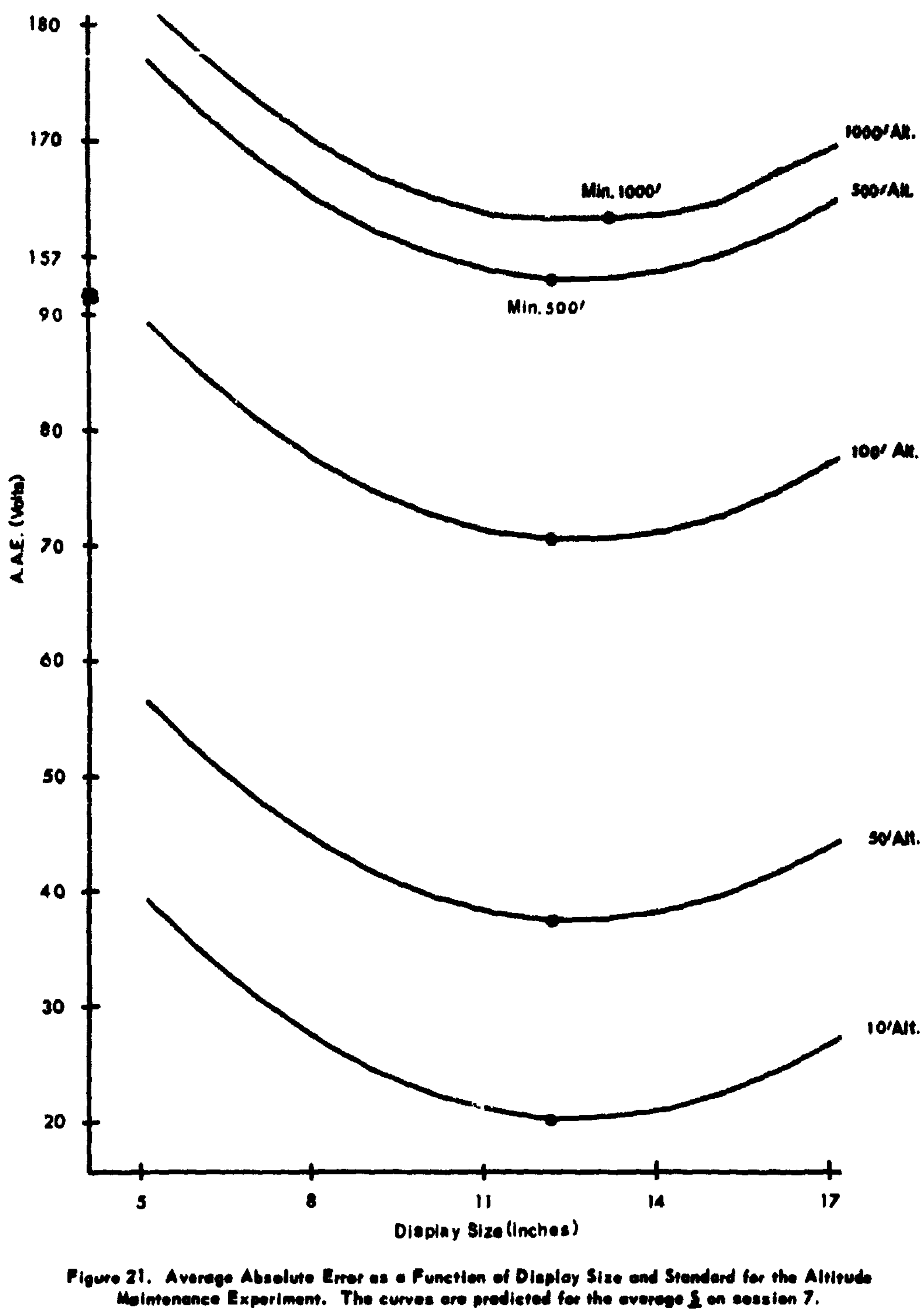


Figure 21. Average Absolute Error as a Function of Display Size and Standard for the Altitude Maintenance Experiment. The curves are predicted for the average S on session 7.

Turning to the effect of display size on AAE, it is seen that intermediate size displays resulted in the least AAE, regardless of the magnitude of the standard. The minimum point of each of the error curves is identified in figure 21. It can be seen that a 12-inch display proved to be optimal for all standards except 1000 feet, in which case a 13-inch display proved best. One other finding which should be noted is that although a Display Size x Standard interaction term was chosen as a predictor of AAE, the fact that all the curves in figure 21 appear to be very nearly parallel indicates that the relative effect of this interaction term was quite small.

Figure 22 shows predicted RMS, for the average S on session 7, as a function of display size and altitude standard. Again, it is seen that error tends to increase as a function of the magnitude of the standard and that error is least for intermediate size displays. Yet several distinct differences between the curves in figure 21 and 22 are apparent. The most obvious difference is the overlapping of the curves for the 500- and 1000-foot standards. The Display Size x Standard interaction evidentally has a more pronounced effect on RMS error than upon AAE. This interaction effect, however, appears to be limited to the two upper standards as the remaining three curves appear to be quite parallel to one another.

The minimum points for the curves in figure 22 are somewhat different from the minimum points of the corresponding curves shown in figure 21. That is, choice of the optimum display size for minimizing judgment error depends upon whether judgment error is indexed by RMS error or AAE. The optimum display size for minimizing AAE was 12 inches for the lower four standards and 13 inches for the 1000-foot standard (see figure 21). Figure 22 shows that the optimum display size for minimizing RMS error was 11 inches for the lower three standards and 9 inches and 8 inches for the 500- and 1000-foot standards.

Another difference in the findings for AAE and RMS error becomes apparent if one considers the relative loss in decoding accuracy that can be expected if a non-optimum display size is used. Figure 21 shows that this loss is generally dependent upon the amount of divergence from the optimum but that relatively smaller losses result from increasing display size than from decreasing it. Figure 22 shows that up to 100-feet altitude, increases and decreases of display size from optimum result in approximately equivalent losses. However, at 500 feet and above, considerably less loss results from using display sizes smaller than optimum.

The RMS error data in figure 22 can be used to derive estimates of the absolute accuracy with which one can expect an average S to maintain a given altitude. Examining judgment accuracy for the optimum display size, it is seen that RMS error varies from 1.1 feet at the 10-foot standard to 35.2 at 1000-feet. Since RMS error is essentially an estimate of the standard deviation of error about the standard, one can use the data

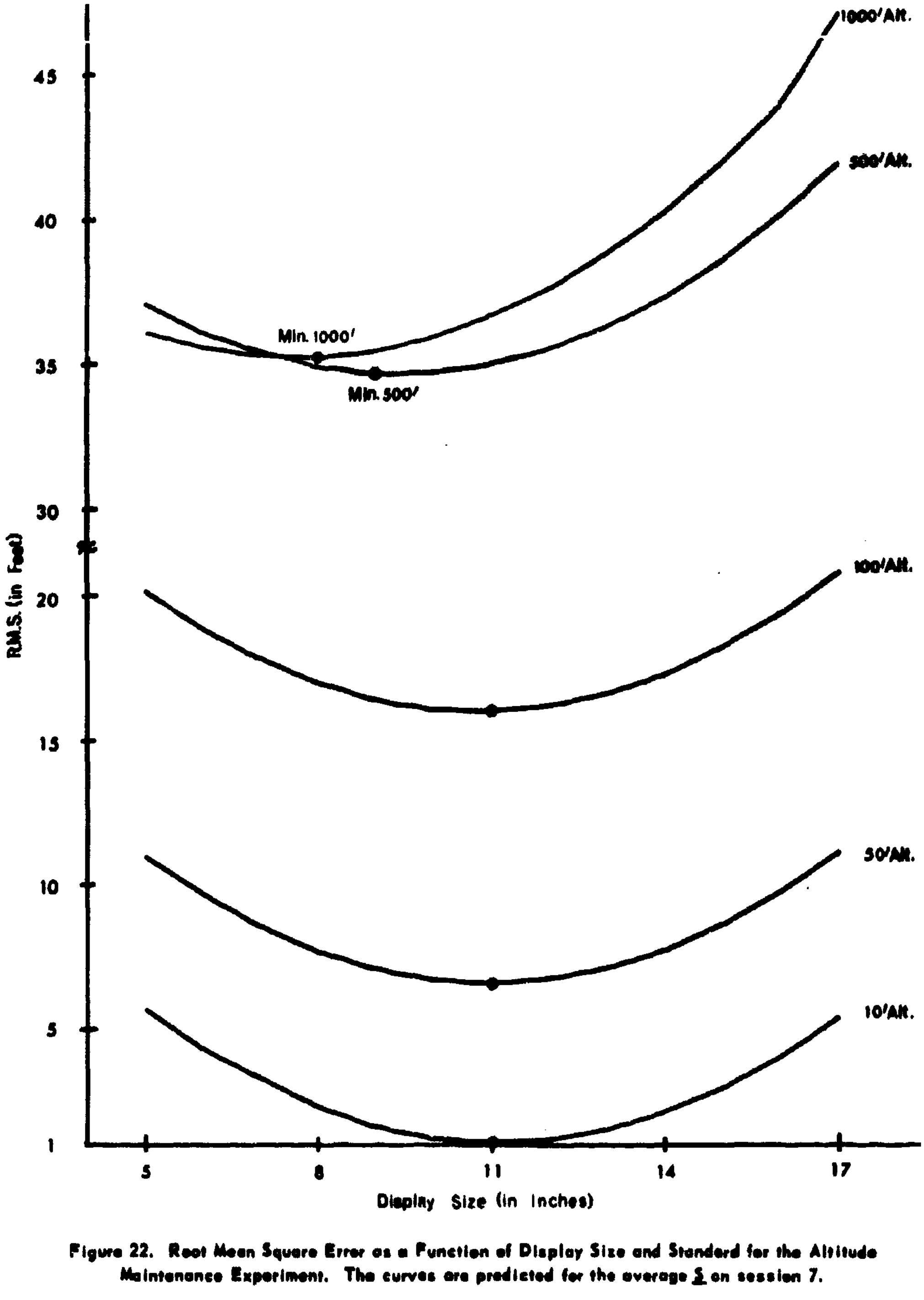


Figure 22. Root Mean Square Error as a Function of Display Size and Standard for the Altitude Maintenance Experiment. The curves are predicted for the average S on session 7.

to define a range about the standard within which the average S should maintain altitude when attempting to maintain that standard.

Discussion

The prediction that the ability to detect and null deviations of altitude from an assigned standard would increase as a function of display size was not totally confirmed by the results of this experiment. In an attempt to account for the curious U-shaped relationship found to exist between display size and judgment error, Ss were questioned in detail as to the cues which they attended to for each display size/standard combination. Although the cues were found to vary widely from standard to standard, they were found to be constant across displays. Thus the U-shaped relationship between display size and judgment accuracy cannot be attributed to the use of different cues for different displays. Furthermore, theoretical consideration of the types of cues that were utilized suggested that larger displays should be more effective. Therefore, it must be stated that no satisfactory explanation is available at this time to explain the particular relationship found.

A prime purpose of this study was to identify the optimum display size for the performance of altitude maintenance. It was found, however, that what is to be considered the optimum display size depends not only upon the magnitude of the altitude standard but also upon whether AAE or RMS error is selected as the criterion of proficiency. It is therefore necessary to consider the relative importance of judgment accuracy at various altitudes and to consider the relative merits of AAE and RMS as indices of proficiency before a final recommendation can be made.

The first problem is simple. Judgment accuracy is obviously most critical at low altitudes so, other things being equal, it seems reasonable to recommend the display size that proves most effective at the lowest altitude.

Some insight into the relative merits of AAE and RMS error as indices of proficiency can be gained if one considers the manner in which they are derived. If it is considered that AAE is based upon the integral of absolute error and RMS error is based upon the integral of squared error, it is clear that RMS error would tend to weight extreme deviations relatively more heavily than AAE. Considering the negative pay-off associated with extreme deviations of altitude from the standard, this attribute of RMS error would appear to be highly desirable. On the basis of the data obtained in this experiment and the rationale presented above, 11 inches is recommended as the optimum size display for altitude maintenance.

Absolute Judgment of Altitude

The purpose of this experiment was to obtain estimates of the accuracy with which the magnitude of altitude can be judged with the basic VCAD, and

to determine whether altitude judgment accuracy varies as a function of display size.

<u>Subjects</u>

Eight laboratory personnel served as <u>S</u>s in this experiment. Four <u>S</u>s were totally familiar with the VCAD, having had served as <u>S</u>s in two previous experiments--including the altitude maintenance experiment. The remaining 4 <u>S</u>s were selected on the basis of having had no prior experience with the VCAD whatsoever.

<u>Apparatus</u>

With two exceptions, the apparatus used in this experiment was identical to that used in the altitude maintenance experiment. Unlike the altitude maintenance study, no forcing function was introduced into the altitude channel in the present experiment. Also, the CADG was modified to provide a numerical read-out of altitude (in feet) in the upper right-hand corner of the display. The experimenter was provided a switch with which to turn these numerics off or on at will.

<u>Procedure</u>

Before beginning the experiment <u>S</u>s were instructed that their task would be to sequentially set the display at one of a number of prescribed altitudes by appropriately manipulating the joystick. Subjects were then given the opportunity to manipulate the stick and observe the resulting changes in the display. After <u>S</u>s reported that they had gotten the "feel" of the control, they were told that they would be shown each of the altitude standards and that they should try to remember the appearance of the display when at each of these altitude standards. Altitude was then sequentially set at 10-, 15-, 20-, 30-, 40-, 50-, 70-, 100-, 250-, 500-, 750-, and 1000-feet, allowing <u>S</u> to view each standard for a period of 30 seconds. Subjects were told that on the first day of their participation they would not be informed of how accurate their settings were, but that on subsequent days numerics would appear on the display--immediately after each setting--which would inform them of the actual altitude of their setting. They were instructed to attend carefully to this feedback information in order to improve the accuracy of subsequent settings.

At the beginning of each trial the experimenter verbally informed <u>S</u> of the standard at which he was to set altitude, e.g., "... your next standard is 750 feet." After <u>S</u> had signaled completion of his setting, the experimenter placed the CADG in a "hold" condition which locked altitude at its current level, threw the switch which activated the altitude numerics, and recorded the altitude of the setting from the experimenter's display monitor. During the "no feedback" trials <u>S</u>s were instructed to look away during the period that the numerics were presented; thereafter, <u>S</u>s were instructed to attend to the numerics and the display configuration

during the inter-trial interval which lasted for approximately 30 seconds.

Twelve altitude standards and four display sizes were investigated, giving a total of 48 experimental conditions. In addition, Ss were tested on one no-feedback session prior to the experiment proper in order to clearly establish the initial level of skill prior to any learning.

Two settings were made for each altitude standard during each daily sesson. The altitude standards were presented in random order with the limitation that the same standard could not be presented on two successive trials. All trials within a daily session were performed on the same size display.

The order of presentation of the 4 displays is illustrated in table 5. This design will be recognized as being composed of two augmented Latin Squares, the characteristics of which were discussed in the "Procedures" section of the previous experiment. The rows adjacent to the S number identifies the order with which that S received the 4 display sizes. Subjects 1, 2, 3, and 4 represent the experienced population and Ss 5, 6, 7, and 8 were non-experienced. Although all 4 display sizes were represented in the no-feedback condition (session zero), it should be noted that the Display Size variable is confounded with the Subjects variable. Because the only objective of session zero was to obtain a valid estimate of the initial level of skill, independent of display size, this confounding was considered unimportant.

Results

Two performance measures were computed from the raw data, viz., AAE and AE. The results of the multiple regression analysis performed on these two measures are summarized in figure 23. Examination of the predicted AE data, an index of judgment bias, showed that judgment biases occurred early in training and were specific to individual Ss rather than being consistent across Ss. But by the end of training these response biases had essentially disappeared. Since S specific response biases which occur only early in learning hold no particular interest, AE will not be discussed further and the reader may consider judgment errors to be symmetrically distributed about the standard by the end of training.

Figure 23 shows that a total of 24 terms were selected for the AAE prediction equation. An F-ratio of 22.2 shows that a statistically significant proportion of the total variance is accounted for by the regression equation (23-percent as per R^2).

As has been true previously, the variable of primary interest in this experiment was Display Size. Figure 23 shows that Display Size was not selected by the regression analysis as contributing to the predictable

Table 5. *Order of Presentation of Display Sizes During Absolute Altitude Judgment Experiment*

	No Feedback	Feedback							
Subjects / Sessions:	0	1	2	3	4	5	6	7	8
S_1 & S_5	17*	17	5	8	14	5	14	17	8
S_2 & S_6	5	5	14	17	8	14	8	5	17
S_3 & S_7	8	8	17	14	5	17	5	8	14
S_4 & S_8	14	14	8	5	17	8	17	14	5

* Display size in inches.

	Display Size	Standard	Sessions	Subjects	Experience	Display Size x Standard	Standard x Sessions	Subjects x Sessions	Standard x Experience	Display Size x Experience	Total Predictors	Total Data Points	F – ratio	p	$\bar{R}^2$	$\bar{R}$
AAE	0	10	1	5	1	1	2	1	1	1	24	1728	22.2	$\leq .005$	.23	.48
AE	1	4	1	4	1	2	5	1	3	0	22	1728	8.8	$\leq .005$	.10	.32

Figure 23. Summary of Results of Multiple Regression Analysis for Absolute Altitude Judgment Experiment.

variance, a finding which indicates that display size had no overall effect on altitude judgment accuracy. However, the finding that a Display Size x Standard interaction term was selected by the analysis indicates that display size did have some slight effect on judgment accuracy but that this effect varied as a function of the altitude standard. Two other variables, whose selection represents no surprise, are Standard and Experience. Although it appears obvious that altitude judgments would vary as a function of the magnitude of the altitude standard and the amount of S's prior experience with the VCAD, it was of interest to determine the relative magnitude of these effects.

The nature of the effects of the variables of primary interest is shown in figures 24 and 25. These figures show AAE as a function of display size and magnitude of the altitude standard. The curves represent predictions for the average inexperienced S (figure 24) and the average experienced S (figure 25) on session 7. Both curves are drawn to approximately the same scale in order to facilitate comparison.

Comparison of figures 24 and 25 will show that altitude judgment error was clearly greater for the inexperienced S. Better than a threefold difference in error magnitude is noted for most altitude standards. Dispite the large difference in the magnitude of judgment error, it can be seen that the relationship among the various curves within the two figures is highly similar. For example, rank ordering the altitude standards on the basis of error shows that the same ranking is obtained for both S populations with one exception. The ranking of the 50- and 500-foot standards for the inexperienced Ss is reversed from that found for experienced Ss.

Of considerably more interest is the finding that judgment error is not a simple increasing function of the magnitude of the altitude standard. The discrepancies from expectation occur at the 50-, 250-, and 750-foot standards. These differences in the relative difficulty of the various altitude standards were found to be a function of the types of altitude cues utilized by the Ss. This topic will be given further consideration in the Discussion section.

The effect of the Display Size x Standard interaction term is apparent in both figures 24 and 25. It is seen that when judging lower altitudes, display size has essentially no effect on judgment accuracy. However, at altitudes above 250 feet, judgment accuracy tends to increase as a function of display size, and the magnitude of this effect tends to increase as a function of altitude.

Referencing the scale on the right-hand ordinate of figures 24 and 25 will enable the reader to interpret the various curves in terms of the standard deviation of error about the altitude standard. As was illustrated in the discussion of previous experiments, these data can be used to define

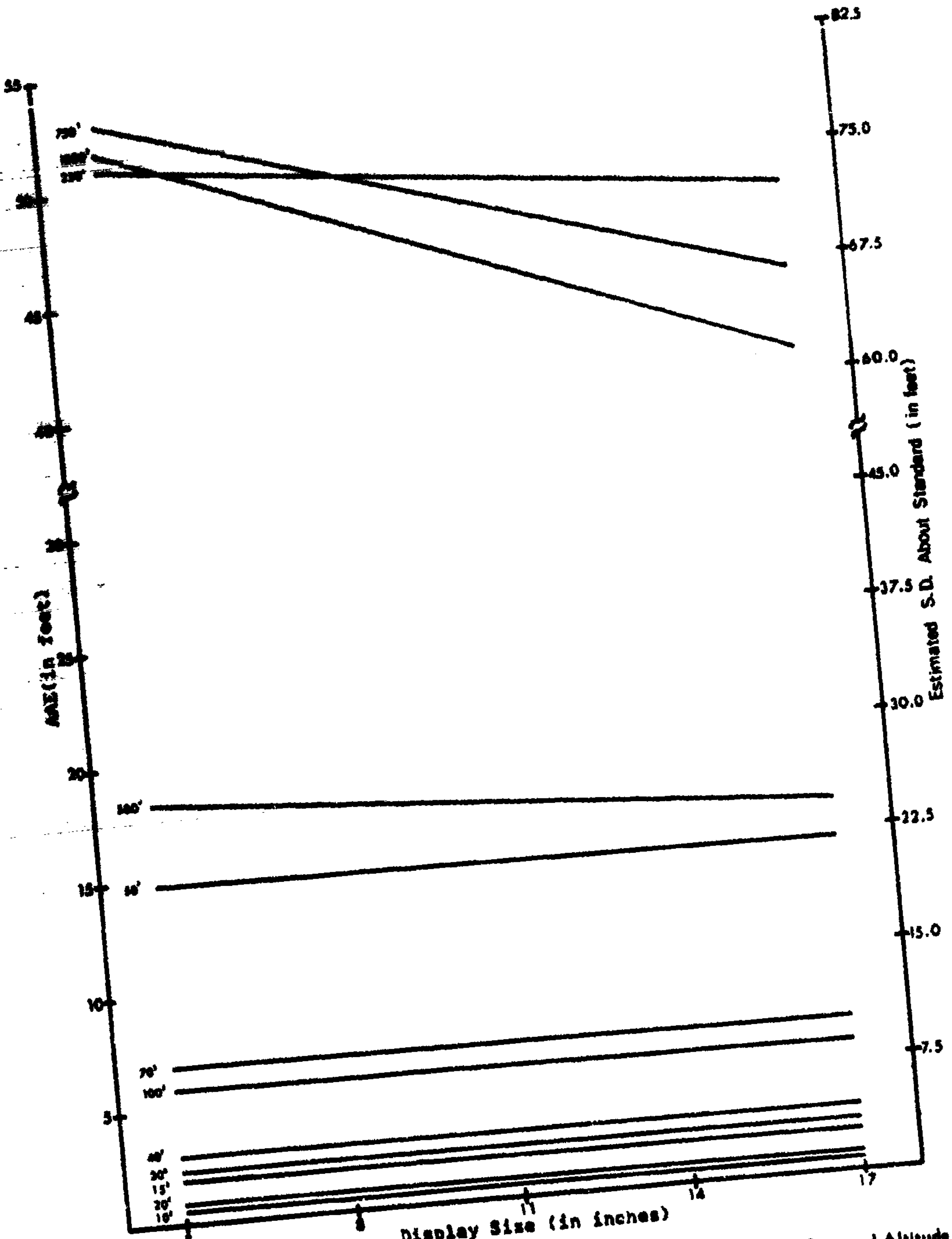


Figure 24. Average Absolute Error in Altitude Settings as a Function of Display Size and Altitude Standard. Predicted data for the average "inexperienced" subject on session 7.

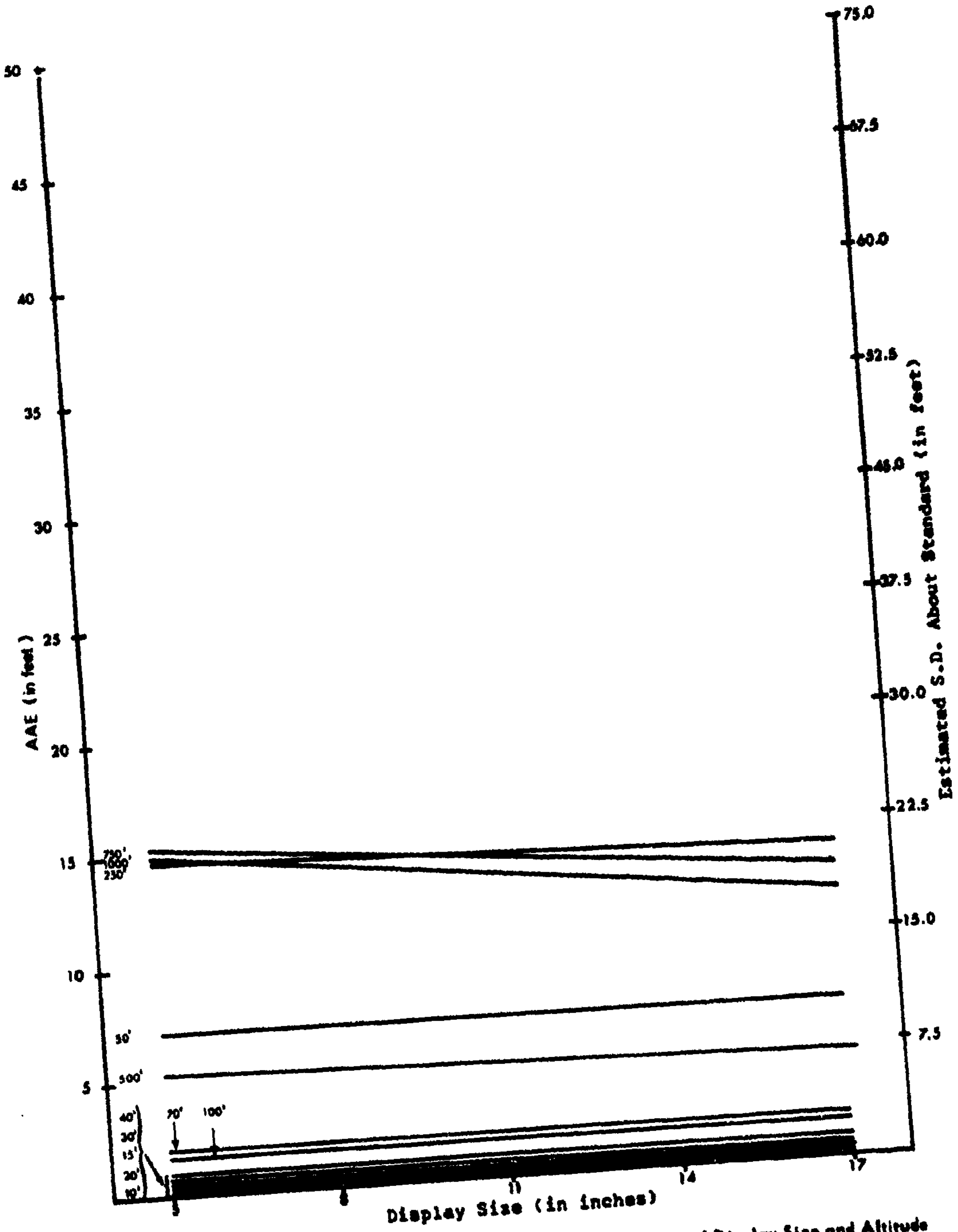


Figure 25. Average Absolute Error in Altitude Settings as a Function of Display Size and Altitude Standard. Prediction for the average "experienced" subject on session 7.

a range about the standard within which a proportion of the average S's judgments could be expected to fall. The width of this range depends, of course, upon the proportion of the population of judgments one desires to encompass within the range.

Discussion

The lack of an overall effect of display size on decoding accuracy is in accord with prior predictions, although the finding of a Display Size x Standard interaction was unexpected. It was originally assumed that the accuracy of altitude judgments would be a function of (a) the absolute amount of change in the altitude cues that results from an increment in altitude--the "sensitivity" of cues to changes in altitude being greater for larger displays and (b) S's ability to remember the location of the standard. Although consideration of the sensitivity of the various cues to changes in altitude would suggest that more accurate judgments would be made with larger displays, it was predicted that the amount of variability due to inaccuracy in recalling the position of the standard would be so large that the effects of display size would be totally masked. If these assumptions are valid, the findings of this experiment suggest that the altitude standards can be more accurately recalled and reproduced than was originally supposed. Consideration of the absolute accuracy with which the altitude settings were made tends to support this conclusion.

A difference between the judgment accuracy of experienced and inexperienced Ss was expected but the size of this difference--even after 7 training sessions--was surprising. The experienced S had been exposed to only the 10-, 50-, 100-, 500-, and 1000-foot standards during the altitude maintenance experiment and they were not specifically instructed to learn to associate quantitative values with these standards. Moreover, their initial proficiency showed that they had not done so. The probable cause of their superior performance was that their prior experience provided them the opportunity to identify a larger number of prospectively useful altitude cues and to differentiate those that were most useful. This knowledge, in turn, enabled the experienced S to better utilize the feedback information.

Examination of the possible causes of the differences in judgment error for the various altitude standards revealed several interesting facts regarding the utilization of altitude cues. It was found that at altitudes of 40-feet and below, no texture breakpoint was available on the display. At these altitudes Ss reported using either size of ground texture or the angle of convergence of perspective lines as cues to altitude. The data in figures 24 and 25 indicate that these were effective cues at these altitudes.

The first order breakpoint first began to emerge clearly at an altitude of 50 feet and was visible up through 750 feet altitude. Most

Ss reported that between 50 and 750 feet altitude they referenced the position of the texture breakpoint relative to either the horizon line or the bottom edge of the display in making their judgments. If it is noted that at low altitudes the position of the breakpoint changes very little as a function of altitude, it follows that breakpoint position is not a sensitive cue to altitude at very low altitudes. The disproportionately large error associated with the 50- and 70-foot standards suggests that it was a mistake for S to use breakpoint position as a cue at these altitudes. Judgment error at these altitudes would probably be considerably reduced if Ss were instructed to attend to either size of ground texture or convergence of perspective lines as their cues to altitude.

Another interesting finding was that judgment error tended to be a function of the proximity of the breakpoint to the bottom edge of the screen. For example, at 250-feet altitude, the breakpoint was located about two-fifths of the distance between the horizon line and the bottom edge of the screen. Judgment errors were very large under this condition. At 500-feet altitude, the breakpoint was located approximately seven-tenths of the distance between the horizon line and the bottom edge of the display and error was found to be even less than at 250 feet. This finding suggests that judgment accuracy would probably improve if a supplementary reference mark were positioned about half the distance between the horizon line and the bottom edge of the display.

PITCH ANGLE EXPERIMENTS

In an earlier discussion it was pointed out that the amount of change in the position of the horizon line that results from an increment in pitch angle is a function of not only pitch angle, but display size and display viewing angle as well. The relationship among these variables was illustrated in figure 5. It is shown in figure 5 that when display viewing angle is manipulated so as to maintain a 1:1 correspondence between the display and the real world, the amount of change in the position of the horizon line that results from a given increment in pitch angle is the same, regardless of the size of display being used. Conversely, if display viewing angle is held constant while display size is varied, the amount of horizon line displacement per pitch increment increases as a direct function of display size. It can thus be seen that larger displays provide a more sensitive index of change in pitch angle than smaller displays - assuming display viewing angle remains constant. However, increasing display size also makes the display less "real-world like" so far as pitch angle is concerned. The larger the display the less the displacement of the displayed horizon line corresponds with the displacement of the "real-world" horizon line, given a fixed pitch increment.

This fact necessitates the modification of research objectives from those described for the roll angle and altitude experiments. Previously, the influcence of display size on both standard maintenance and absolute judgment tasks was investigated. Although measuring a S's ability to maintain a given standard pitch angle as a function of display size is a legitimate procedure, once a S is asked to ascribe numerical values to a given pitch angle (absolute judgment) the task becomes impossibly confusing. If a S were to judge the position of the horizon line according to a "real-world" criterion, his error would automatically increase as a function of display size. If a S were asked to judge roll angle according to the CADG read-out, he would be asked to learn a totally different scale for each display size and his proficiency would be inversely related to his ability to ascribe numerical values to "real-world" pitch angles. Either way, measuring absolute pitch angle judgments as a function of display size would produce totally meaningless results.

The objective of the first experiment reported in this section was to define the nature of the relationship between display size and accuracy with which a standard pitch angle can be maintained, and to obtain a reliable estimate of the accuracy with which the maintenance task can be performed. The objective of the second experiment was to obtain a reliable estimate of absolute roll angle judgment accuracy for a single display size.

Pitch Angle Maintenance

Subjects

Four laboratory personnel served as Ss in this experiment. The Ss had neither previous flight experience nor previous experience with the VCAD.

Apparatus

With the exception of the following features the apparatus was the same as that described in the "General Apparatus" section of this report.

1. Four display sizes were available for investigation (5, 8, 14, and 17 inches).

2. Sine wave forcing functions were introduced into both the pitch and roll channels. The frequency and amplitude of the forcing function introduced into the pitch channel were 2.6 cpm and ±15 degrees. This forcing function caused pitch angle to vary symmetrically about whichever pitch angle standard was in use at the time the forcing function was introduced. The forcing function introduced into the roll channel had a frequency of 2.0 cpm and an amplitude of ±12 degrees. Roll angle varied about zero degrees throughout the experiment.

3. The S was provided the capability of nulling error in roll angle by feeding the joystick output into the pitch channel. Displacement of the stick forward of the null position introduced a negative rate causing the simulated aircraft to pitch down. Displacement in the opposite direction introduced a positive rate causing the simulated aircraft to pitch up. No control of roll angle was provided.

4. Circuitry was developed to enable the experimenter to set pitch angle at plus or minus 75, 45, 30, 15, or 0 degrees by manipulating a potentiometer located on the analog computer. These nine conditions represented the roll angle standards investigated. Also, the equipment was configured such that pitch angle would automatically return to the standard pitch angle at the termination of each trial.

5. The performance measures obtained included integrated absolute error, integrated error squared, and integrated error with respect to sign.

6. Altitude was held constant at 500 feet and forward ground velocity was maintained at 150 knots. No lateral movement was introduced.

Procedure

Prior to the first experimental trials, Ss were instructed that their task was to maintain an assigned pitch angle throughout a two-minute trial by manipulating the control-stick in the longitudinal plane. Subjects were then allowed to manipulate the control in order to become familiar with its dynamics and with the display/control relationship. Before performing on each of the standards, Ss were permitted to view the standard until they felt sure they could recognize its appearance. Thereafter, Ss were instructed to devote the inter-trial interval attempting to better learn the appearance of the standard. Subjects were given verbal signals to indicate the start and the termination of each trial.

Each of the nine roll angle standards were paired with each of the four displays, giving a total of 36 experimental conditions. Subjects were given three consecutive two-minute trials per daily session on each of the roll angle standards except for the zero-degree standard in which case they were given two temporally separate sets of three trials each. All the 30 trials within a daily session were performed on the same size display.

The nine roll angle standards were randomized within sessions. The order with which each S received the different display sizes was established by using augmented Latin Squares as shown in table 6. The features of this design were discussed in conjunction with previous experiments.

Table 6. Order of Presentation of Display Sizes During Pitch Angle Maintenance Experiment

Subjects	Sessions 1	2	3	4	5	6	7	8
S_1	14*	17	8	5	14	17	8	5
S_2	17	8	5	14	17	8	5	14
S_3	8	5	14	17	8	5	14	17
S_4	5	14	17	8	5	14	17	8

* Display size in inches.

Results

The three performance measures computed from the raw data were AAE, RMS error, and the standard deviation about the mean of the error distribution. The latter two of these measures were found to be very highly correlated ($r = .998$). As was discussed previously, this finding indicates that the mean of the error distribution is zero (no judgment bias) and that errors must therefore be symmetrically distributed about the pitch angle standard. Since both measures are essentially identical, only RMS was submitted to analysis.

The results of the multiple regression analyses for AAE and RMS error are summarized in figure 26. It is seen that a total of 19 predictor variables were chosen for each measure. The composite of these predictor variables was found to account for a highly significant proportion of the total variance. The proportion of the total variance accounted for by the regression equation (as indicated by $\bar{R}^2$) can be seen to be 73 percent and 72 percent for AAE and RMS, respectively.

Six different sets of predictor variables were chosen by the multiple regression analysis. Of these the Display Size variable and the Standards variable hold the most interest. The effects of these variables are discussed in the following pages.

Figure 27 shows the relationship between AAE (in volts) and display size. These data represent predictions for the average S maintaining the zero-degree standard on session 7. Since there was no interaction between display size and standard, the shape of this curve (but not the magnitude of error) can be considered representative for all standards. It can be seen that AAE was found to be least for intermediate size displays. The minimum point on this curve was found to correspond with a 12.25-inch display.

RMS error is shown in figure 28 as a function of display size and magnitude of the pitch angle standard. Because of the close proximity of the curves for the various display sizes, only the curves for the worst and the best cases are shown. The curve for the 5-inch display represents the worst case and the curve for the 12.25-inch display, the best. The curves for all other display sizes fall within these bounds. The most striking effect in figure 28 is the finding that the minus 75-degree standard was clearly more difficult than the remaining standards. The probable reason for this finding will be discussed in the following section. Otherwise, there appears to be little difference in RMS error among the remaining standards. Of considerable interest is the absolute magnitude of the RMS error for the various standards. Considering the 12.25-inch display, RMS error is seen to be less than one degree for all standards except the negative 75-degree standard.

	Display Size	Standard	Sessions	Subjects	Standard x Sessions	Subjects x Sessions	Total Predictors	Total Data Points	F-ratio	p	$\hat{R}^2$	$\hat{R}$
AAE	2	8	2	2	3	2	19	960	139.3	$\leq .005$	.734	.857
RMS	2	8	2	1	5	1	19	960	132.4	$\leq .005$	.724	.851

Figure 26. Summary of Results of Multiple-Regression Analysis for Pitch Angle Maintenance Experiment.

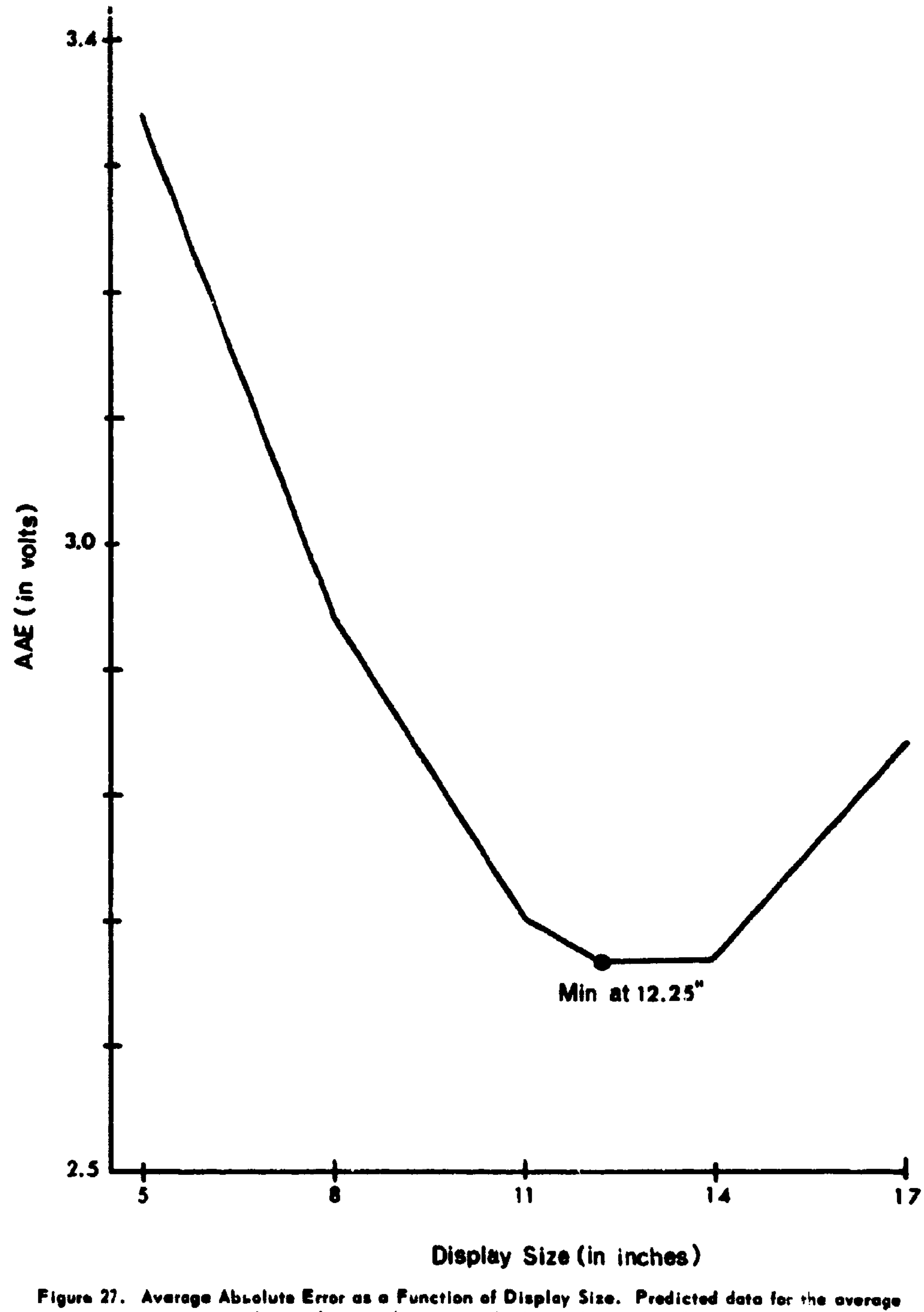


Figure 27. Average Absolute Error as a Function of Display Size. Predicted data for the average subject, the zero-degree standard, and for session 7.

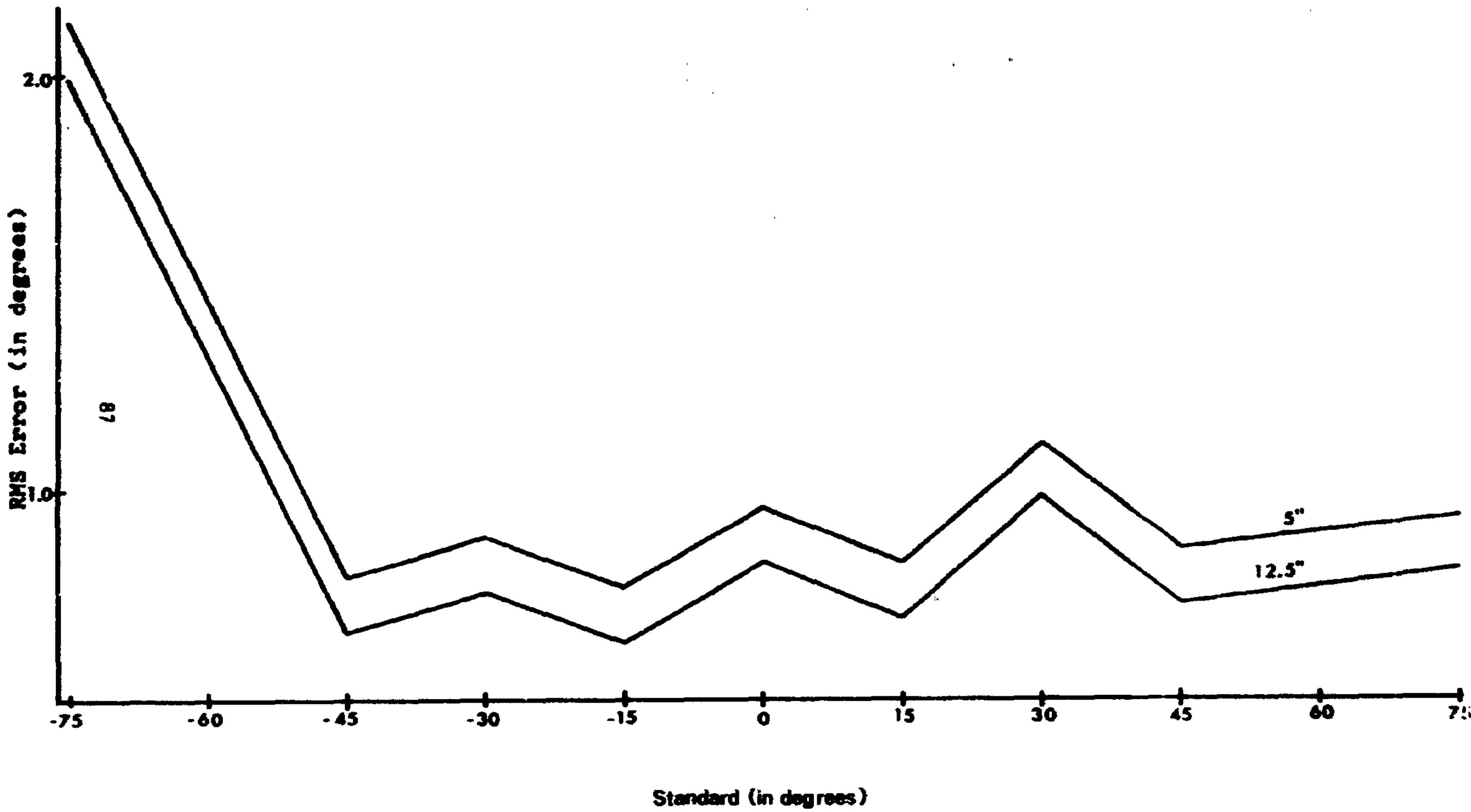


Figure 28. Root Mean Square Error as a Function of Display Size and Magnitude of Standard. Predictions for the average subject on session 7.

Since RMS error is an estimate of the standard deviation of error about the standard, it can be stated with a high degree of confidence that the average S could maintain pitch angle to within 3 degrees of any pitch angle standard (from negative 45 to positive 75 degrees) a very high proportion of the time.

Discussion

The prediction that the accuracy of pitch angle maintenance would increase as a function of display size was only partially confirmed by these data. Again, no good explanation is available at this time to account for the curious "U" shaped relationship between maintenance judgment accuracy and display size. Although the effect of display size was significant from a statistical view, display size over the range investigated has little practical import. The largest RMS error difference between the worst and the optimum display was only about 10-minutes of arc.

The relative difficulty of the negative 75 degree standard was found to be a function of the ineffective pitch angle cues available at that condition. It was found that either the horizon line or a breakpoint line was visible for all standards except positive and negative 75 degrees. In the case of a positive 75 degree standard, sky texture elements and the Zenith Marker were visible and provided stable references for pitch. When pitched down 75 degrees, neither the horizon line or a breakpoint line could be seen, thus forcing Ss to reference the angle of convergence of perspective lines in judging pitch. The magnitude of the error associated with this condition makes it apparent that perspective lines do not provide an effective cue for pitch angle. In comparing the magnitude of error for each of the pitch angle standards with the strategies used, it was generally found that error was inversely related to the proximity of a stable reference (horizon line, breakpoint line, sky texture elements, Zenith Marker) to the upper or lower edge of the display.

Absolute Pitch Angle Judgments

The objective of the present experiment was to obtain a reliable estimate of the accuracy with which a S can be expected to judge the absolute magnitude of pitch angle. Because of the accuracy with which Ss were able to maintain the various pitch angles in the previous experiments and because some of the cues utilized would not be available in a flight situation, it was decided to measure absolute pitch angle judgments under conditions which provided a S with fewer stable cues. The introduction of forcing functions in the roll and the altitude channels forced a S to abandon some of his "alignment" cues and to depend more upon his memory of the appearance of a condition than upon his ability to align some stable features of the display with the edge of the display. At negative pitch angles Ss are forced to use angle of

convergence of perspective lines as cues and since the sensitivity of this cue varies as a function of altitude it was also of interest to determine the effect of altitude on the accuracy of judging large negative pitch angles. Since the sky plane is always seen at a constant distance, changing altitude could have no effect on the judgment of large positive pitch angles.

Subjects

Three laboratory personnel were used as Ss in this experiment. All three Ss had served previously in an experiment concerned with absolute pitch angle judgments. That experiment was scrapped because the cues utilized were not considered representative of those available in a flight situation. All Ss can therefore be considered highly trained upon participating in this experiment.

Apparatus

The apparatus used in the pitch angle maintenance experiment was modified only slightly for use in the present experiment. As mentioned previously, forcing functions were introduced into both the roll and the altitude channels. Altitude was driven off its initial level by ±50 feet and roll angle traveled through an excusion of ±6 degrees about zero. The frequency of the forcing functions was 2.0 cpm for altitude and 2.3 cpm for roll angle. A direct read-out of momentary pitch angle in degrees was provided on a digital voltmeter located on the analog computer. An 8-inch display was used throughout the experiment.

Procedure

Nineteen pitch angle standards, varying from negative 90 degrees to positive 90 degrees in 10-degree increments, were investigated. The accuracy with which each of these standards could be judged was measured at each of three altitudes (100, 500, and 1000 feet). Each of the standards were judged once during a daily session. The order of presentation of the standards was randomized within sessions and all judgments within a session were performed at the same altitude. Altitude was counterbalanced across sessions as shown in table 7. This design ensures that each S receives each altitude with equal frequency, that the altitudes are presented in all possible orders, and that each altitude is represented at each stage in practice.

The experimental procedures were the same as used in the absolute altitude and absolute roll angle judgment experiments and will not be repeated here. Because all Ss had participated in a prior experiment concerned with absolute pitch angle judgments it was unnecessary to re-indoctrinate them regarding the procedures. Feedback was given immediately after each judgment by verbally informing the S of the actual magnitude of his setting in degrees.

Table 7. Order of Presentation of Altitudes During Absolute Pitch Angle Experiment

Subjects	Sessions 1	2	3	4	5	6	7	8	9
S_1	1000*	500	100	100	1000	500	500	100	1000
S_2	100	1000	500	500	100	1000	1000	500	100
S_3	500	100	1000	1000	500	100	100	1000	500

* Altitude in feet.

Results

The results of the multiple regression analyses of AAE and AE are summarized in figure 29. A total of 30 predictor variables were selected for AAE and 21 for AE. A statistically significant proporation of the total variance was accounted for by both regression equations (Note the F-ratios and the $\overline{R}^2$ values). This study was designed to investigate the effects of both pitch angle standard and altitude on absolute judgment accuracy. Although only the former was selected by the regression analysis, it will be noted that a Standard X Altitude interaction variable was selected indicating that altitude was not entirely without effect.

The nature of the effects of these two variables is illustrated in figure 30 and 31. Figure 30 shows AAE as a function of magnitude of altitude and pitch angle standard. Different altitude curves are not shown for positive pitch angles because the display configuration was essentially independent of altitude at these angles. It can be seen that AAE was found to be relatively stable over the range between negative 40-degrees and positive 90-degrees pitch angle. As pitch angle is decreased below negative 40-degrees, AAE tends to first increase and then decrease at negative 90-degrees. Although the accuracy of pitch angle judgments clearly varies as a function of altitude, the relationship is in no way systematic.

Before the meaning of the AAE scores is discussed further, attention will be turned to figure 31 which shows AE as a function of magnitude of altitude and pitch angle standard, remembering that AE is an index of judgment bias. Figure 31 shows that sizable judgment biases were present at certain standards, and for certain altitudes. The largest biased occurred at the negative 90 degree standard when judging from an altitude of 1000 feet. This AE socre indicates that, on the average, a S tended to adjust pitch to an angle about 10 degrees less than the standard when attempting to set pitch angle to negative 80 degrees. Other deviations from the zero line are interpreted similarly.

The important effect of these judgment biases is to limit the interpretation that can be placed on AAE. Previously, when it was found that errors were distributed symmetrically about the standard, one could take advantage of the equality, SD $\hat{=}$ 1.5 (AAE), in making statements about the judgment accuracy that would be expected from the average S. However, this quantity is not a reliable estimate of the standard deviation of error about the standard when symmetry of error about the standard does not exist.

Discussion

Few general statements can be made regarding the effect of magnitude of the pitch angle standard and altitude on the accuracy of pitch angle

	Standard	Sessions	Subjects	Standard x Sessions	Subjects x Sessions	Standard x Altitude	Total Predictors	Total Data Points	F-ratio	p	$\tilde{R}^2$	$\tilde{R}$
AAE	8	1	1	10	2	8	30	540	16.5	≤.005	.48	.69
AE	8	0	0	5	0	8	21	540	7.7	≤.005	.23	.48

Figure 29. Summary of Results of Multiple Regression Analysis for Absolute Pitch Experiment.

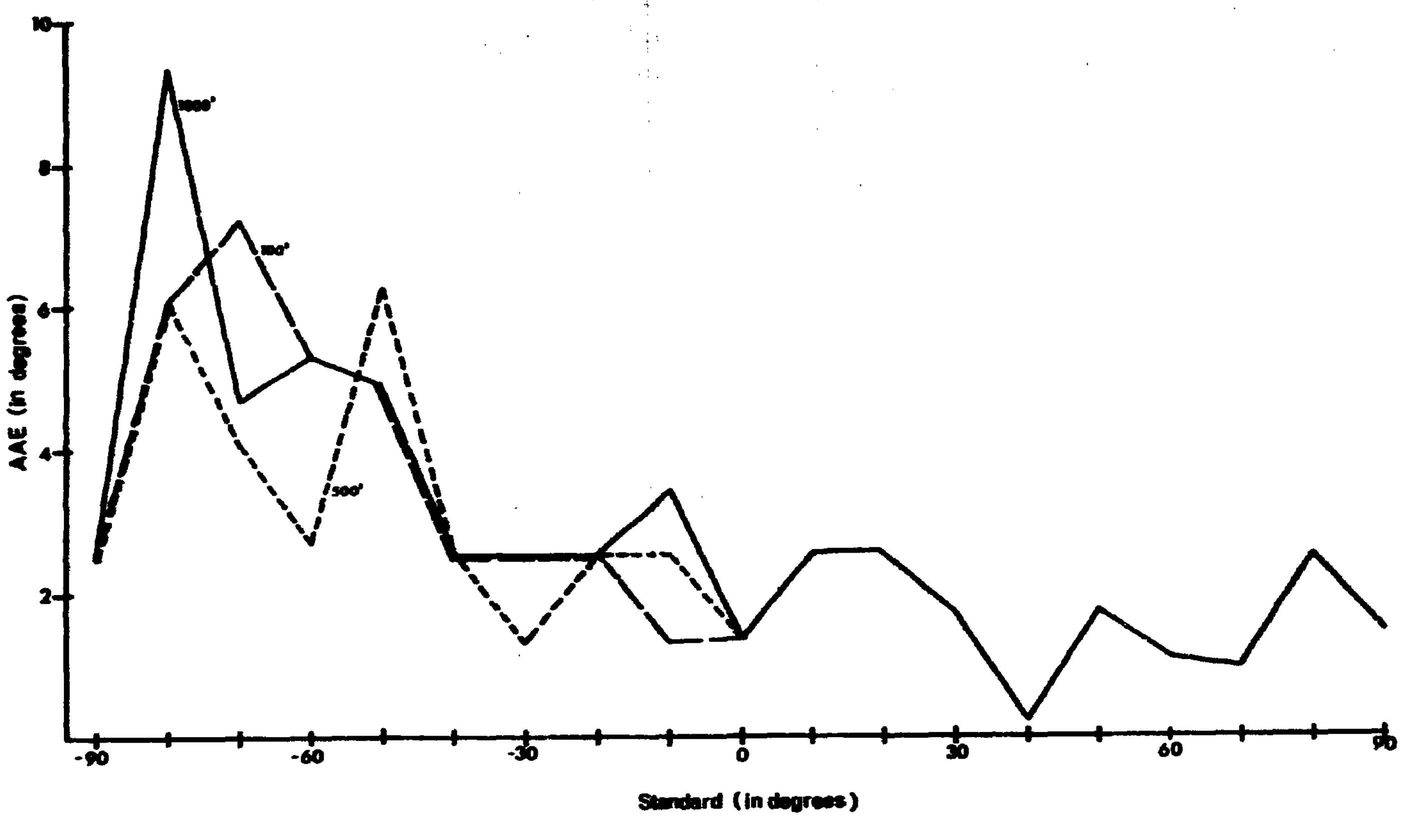


Figure 30. Average Absolute Error of Pitch Angle Judgments as a Function of the Magnitude of Altitude and Pitch Angle Standard. Predicted data for the average subject on session 8.

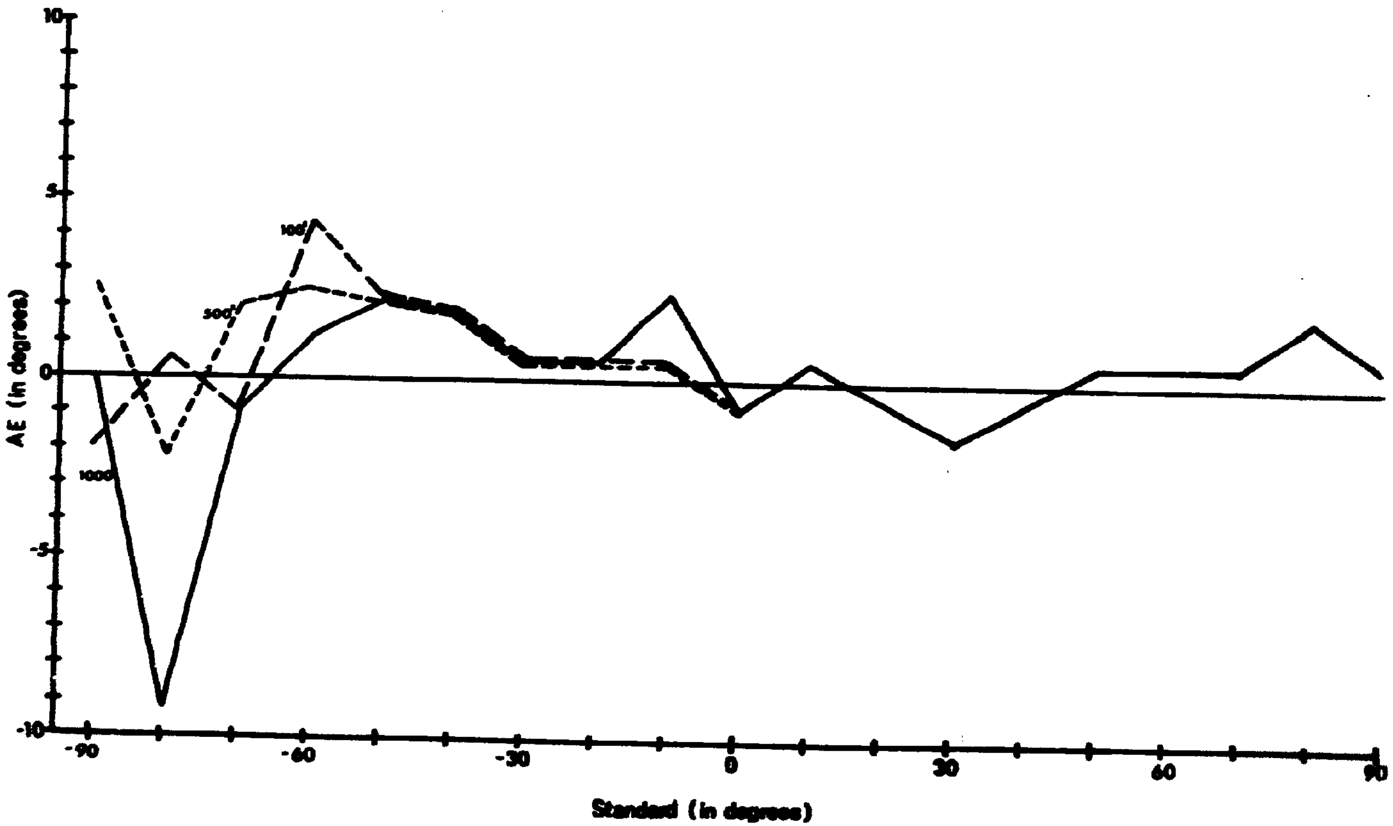


Figure 31. Average Error of Absolute Pitch Angle Judgments as a Function of Altitude and Standard. Predictions for the average subject on session 8.

judgments. Although both altitude and standard affect the accuracy of judging negative pitch angles, this effect does not appear to occur in any predictable way.

In using these data to make predictions as to expected judgment accuracy, it is necessary to exercise considerable caution due to the presence of sizable judgment biases. However, simultaneous reference to figures 30 and 31 allows one to obtain a rough estimate of the magnitude and direction of error for each of the standards.

GENERAL DISCUSSION AND CONCLUSIONS

Display size was found to have an effect on the accuracy of roll angle, pitch angle, and altitude judgments, but the nature of the effect depended upon the type of judgment (absolute judgments vs detection of change) and the magnitude of the standard being judged. The generalization can be made that display size had little effect on the accuracy with which absolute judgments of the three flight parameters can be made, although there was a tendency for judgment accuracy to improve with display size, when Ss were judging altitudes greater than 250 feet.

Display size was consistently found to have an effect on the accuracy with which an assigned value of one of the parameters can be maintained. Intermediate size displays resulted in the most accurate maintenance of altitude and pitch angle regardless of the magnitude of the standard, although the "optimum" size display varied somewhat as a function of the standard. Considering the altitude and the pitch angle studies jointly the optimum display size fell between 8 and 13 inches regardless of the value of the standard. The display most frequently found optimum in these studies was the 11-inch display. In contrast, a rather striking interaction between display size and magnitude of the roll angle standard was found. Larger displays were favored if the standard was less than 18 degrees, whereas, smaller displays were best if the standard was greater than 18 degrees. The display size that appeared to maximize judgment accuracy across all roll angles was the 11 inch display.

When the results of these experiments are considered as a unit, the 11-inch display would appear to be optimum. However, the reader is cautioned against accepting this figure without first considerning the relative priority of the judgments which will be performed with the display and the values of standards that will be judged in a particular flight situation, if a flight situation were encountered in which it was urgent that the pilot be able to identify 1000-feet altitude with a high degree of accuracy, a 17-inch display, rather than an 11-inch display would be considered optimum. One of the prime reasons for defining relationships among variables was to permit the making of such decisions. The reader is also cautioned against generalizing this figure without regard to the distance at which the display will be viewed. Actually, display size

should have been expressed throughout this report in terms of degrees of visual angle subtended by the displays. However, to have done so would have been burdensome to both the writer and the reader. If one intends to generalize the results of these experiments to situations in which viewing distance is not the same as that used here, all display sizes should be converted to units of visual angle in order to make the two siuations equivalent.

To aid in this conversion, a table is provided in appendix A which lists the measured height, width, and diagonal dimensions of the picture shown on commercial CRT tubes varying from 5 inches to 17 inches in increments of one inch. The corresponding visual angle subtended at a 32-inch viewing distance is shown adjacent to each measured dimension.

Generalizations as to the accuracy with which the flight parameters can be judged cannot be made because of the dependency of judgment accuracy upon both display size and the value of the standard being judged. However, the regression equations derived or the figures presented here can be used to estimate the expected judgment accuracy for any specific condition of interest within the range of values investigated. Again a word of caution regarding the interpretation of these results. The accuracy attained in these experiments should not be generalized to complex flight situations. These experiments were concerned with decoding accuracy under near optimal conditions. The introduction of increased perceptual loading or more complex encoding tasks would undoubtedly result in decreased accuracy. These data are to be used as baseline data in determining whether judgments can be made with sufficient accuracy to meet system requirements. If judgments in this near optimum condition are not sufficiently accurate to meet system requirements, one can be sure that the display will require supplementation in a more complex situation. If accuracy decreases as the complexity of the task increases, these baseline data will provide the knowledge that the magnitude of the error is not due to the observer's inability to decode the display but to other factors such as task loading, display/control incompatibility, task complexity, etc.

One of the most valuable products of this series of experiments was the identifcation of the types of cues which Ss tend to use in making various types of judgments. Although these cues are too numerous to list here, it can be stated that nearly all Ss showed a tendency to "externalize" their judgments. That is, Ss tended to search for two stable references on the display and attempted to maintain a constant spatial relationship between these two points. Yet, when these constant references were eliminated, it was found that Ss were capable of learning to judge the parameters accurately using "internal" references (memory presumably). Several cues proved to be unexpectedly powerful, while others proved of less value than was originally presumed. Systematic statements as to the relative value of the various types of cues on the VCAD must, however, await further research.

REFERENCES

1. Cochran, W.G. & Cox, G.M. Experimental Designs (2nd Ed). Wiley, New York, 1957.

2. Cross, K.D. Annual Report - JANAIR Vertical Contact Analog Display Evaluation Program, Naval Missile Center, Feb 1968, (in press)

3. Darlington, R.B., Multiple regression in psychological research and practice. Psychological Bulletin, 1968,69, 161-182.

4. Fedderson, W.E., Experimental evaluation of an eight-inch CRT contact analog presentation. Bell Helicopter Company Report No. D228-421-011 May 1962.

5. Fox, G.A., Hardt, H.D., and Matheny, W.G., Detection of Small Changes in the Size of the Squares in a grid line display. Bell Helicopter Report No. D228-420-002, Feb 1959.

6. Guilford, J.P., Psychometric Methods. McGraw Hill, New York, 1954.

7. Sheffe', H. The Analysis of Variance. Wiley, New York, 1959.

8. Thorndike, R.L. Personnel Selection. Wiley, New York, 1949.

9. Thurstone, L.L. A law of comparative judgment. Psychological Review, 1924, 34, 273-286.

10. Wherry, R.J. A new formula for predicting the shrinkage of the coefficient of multiple correlation. Annals of Mathematical Statistics, 1931, 2, 440-457.

11. Wherry, R.J., Jr. Personal Communication

12. Wilkerson, L.E. & Matheny, W.G., Some factors bearing on size, shape and placement of a vertical analog display presenting hovering information. Bell Helicopter Company Report No. D228-421-006, February, 1961.

13. Williams, P.R., Technical progress report on Universal Contact Analog Display (VCAD) Research. Phase I - Systems Analysis. Norden Report number 1161-R-0011, April 1965.

14. Williams, P.R. & Kronholm, M.B., Technical report on simulation studies of an integrated electronic vertical display. Norden Report number 1161-R-0021, December 1965.

15. Instruction manual for digital contact analog, Volume A - theory, operation and maintenance. General Electric (Contract No. noNR 4057(00)), May 1966.

16. Instruction manual for digital contact analog, Volume B - computer program listings. General Electric (Contract No. noNR 4057(00)). may 1966.

APPENDIX

Table 1a. Conversion Table for Converting Commercially Designated Display Sizes to Visual Angle

Commercial Size Designation (in inches)	Height		Width		Diagonal	
	Measured (in inches)	Visual Angle (in degrees)	Measured (in inches)	Visual Angle (in degrees)	Measured (in inches)	Visual Angle (in degrees)
5	3.4*	6.1**	4.4	8.0	4.9	8.8
6	4.1	7.3	5.3	9.6	5.9	10.3
7	4.7	8.5	6.2	11.0	6.8	12.1
8	5.4	9.5	7.0	12.5	7.8	13.9
9	6.0	10.9	7.8	14.0	8.7	15.5
10	6.7	11.9	8.7	15.5	9.6	17.0
11	7.3	13.0	9.5	17.0	10.5	18.6
12	7.9	14.0	10.3	18.3	11.4	20.2
13	8.5	15.2	11.1	19.7	12.3	22.1
14	9.1	16.2	11.9	21.1	13.2	23.3
15	9.7	17.3	12.7	22.4	14.0	24.6
16	10.3	18.3	13.5	23.7	14.9	26.2
17	10.9	19.4	14.2	25.0	15.7	27.6

* The measured height, width, and diagonal dimension of a given commercially designated display the same for all manufacturers.

** The visual angles were computed for a viewing distance of 32 inches.

$$VA(deg) = 2 \tan^{-1} \frac{\text{Measured inches}}{64''}$$

UNCLASSIFIED
Security Classification

DOCUMENT CONTROL DATA - R & D

(Security classification of title, body of abstract and indexing annotation must be entered when the overall report is classified)

1. ORIGINATING ACTIVITY (Corporate author) Naval Missile Center Point Mugu, California	2a. REPORT SECURITY CLASSIFICATION UNCLASSIFIED 2b. GROUP --

3. REPORT TITLE

JANAIR VERTICAL CONTACT ANALOG DISPLAY EVALUATION PROGRAM; Accuracy of Altitude, Roll Angle, and Pitch Angle Judgments As a Function of Size of Vertical Contact Analog Display

4. DESCRIPTIVE NOTES (Type of report and inclusive dates)

5. AUTHOR(S) (First name, middle initial, last name)

K. D. Cross and A. C. Bittner

6. REPORT DATE 31 January 1969	7a. TOTAL NO. OF PAGES 100	7b. NO. OF REFS 16

8a. CONTRACT OR GRANT NO.	9a. ORIGINATOR'S REPORT NUMBER(S)
b. PROJECT NO. AIRTASK RAV-09M001 (A)	TM-69-2
c. ONR Project Orders 8-0079, 8-0080, 8-0081, 8-0082.	9b. OTHER REPORT NO(S) (Any other numbers that may be assigned this report)
d.	JANAIR Report No. 680713

10. DISTRIBUTION STATEMENT

This document has been approved for public release and sale; its distribution is unlimited.

11. SUPPLEMENTARY NOTES	12. SPONSORING MILITARY ACTIVITY Naval Air Systems Command and Office of Naval Research (Code 461)

13. ABSTRACT

The purpose of this document is to report the first results from a series of experiments establishing and optimizing the utility of the Vertical Contact Analog Display (VCAD). The VCAD concept is that of a computer-generated pictorial display which provides the observer with visual cues as to aircraft orientation. These cues are analogous to those available under contact flight conditions. The specific objectives of this experiment series were: (1) to establish the relationship between size of VCAD display and judgment accuracy for altitude, pitch, and roll estimates, and (2) to obtain estimates of the accuracy with which these quantities may be judged. Four studies were conducted to measure Ss ability to "maintain" a given standard against a forcing function and three studies were conducted to measure Ss ability to recall and utilize an "internalized standard." Both types of experiments were conducted for each of the three flight related quantities. The experimental designs for each of these experiments were fractional replicates with repeated measures. The analytic procedure consisted of the fitting of "response surface" models and the graphical representation of these models. It was concluded that display size had an effect on the accuracy for each of the three flight quantities, but the nature of the effect depended on both the type of judgment task ("maintenance" versus recalled "standard") and the magnitude of the standard being judged. Generally, the effect was more pronounced for the "maintenance" experiments. The results of the experiments as a unit suggest the optimal display traverses a diagonal visual angle of 18.6 degrees for commercial-rectangular displays (e.g., a commercially designated 11-inch display at a 32-inch viewing distance). ( )

DD FORM 1 NOV 65 1473 (PAGE 1)

S/N 0101-807-6801

UNCLASSIFIED
Security Classification

UNCLASSIFIED
Security Classification

14. KEY WORDS	LINK A		LINK B		LINK C	
	ROLE	WT	ROLE	WT	ROLE	WT
Vertical Contact Analog Display (VCAD)						
Display size						
Pitch judgments						
Roll judgments						
Altitude judgments						

DD FORM 1 NOV 65 1473 (BACK)
(PAGE 2)

UNCLASSIFIED
Security Classification